Master thesis for the Master of philosophy of Economics degree

Predicting Domestic Public Expenditures on HIV and AIDS in Low and Middle Income Countries

Identifying Main Determinants

Hanne Gaup Braa

July 2007

Department of Economics

University of Oslo

Preface

This thesis has been written in connection with Joint United Nations Programme on HIV/AIDS (UNAIDS), for the Resource Tracking and Projections Unit of the Evaluation and Monitoring department in Geneva, Switzerland. I would like to thank my supervisor in Geneva, Carlos Avila-Figueroa for inspiring me and sharing his knowledge and expertise of HIV and AIDS. Secondly, I would like to thank co-supervisor at the Department of Economics of the University of Oslo, Professor Erik Biørn, for invaluable help with the analysis and for his precise feedback. I would also like to thank Jose Antonio Izazola for his work with the database and for his inspiring discussions. Finally, I would like to thank Angela Bulgari and Karen Petterson for helpful comments and corrections, Christian Thorkildsen for his help and support, and Kolbjørn Braa for being my source of motivation.

Hanne Gaup Braa
Geneva, July 13th 2007

Abstract

The project of which this thesis is a part of is aimed at developing a comprehensive framework for health and non-health public expenditures to confront the HIV epidemic. Within the framework of a regression model, the main drivers of the domestic public expenditures on HIV and AIDS is identified, and these expenditures is predicted where data on this variable is missing. The model analyzes the factors associated with government disbursements related to HIV/AIDS in low- and middle-income countries, using information from the National AIDS Spending Assessments at the country level. It considers the impact of demographic and non-demographic variables for both health and non-health (i.e. education, advocacy and policy development, the judiciary system and human rights) expenditures. The results are meant to provide essential empirical information on financial resources. Identifying the factors associated with domestic public expenditure is important when examining the countries' ability to take on a greater share of the financial burden to sustain a response to the HIV epidemic.

Abbreviations and Acronyms

AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
CAR	Caribbean
CPI	Corruption Perceptions Index
CRIS	Country Response Information System
CRS	Creditor Reporting System
domexcap	The domestic public expenditures on HIV and AIDS per capita
EA	East Asia
EECA	Eastern Europe and Central Asia
GDP	Gross Domestic Product
GFATM	Global Fund to Fight AIDS Tuberculosis and Malaria
GNI	Gross National Income
HIV	Human Immunodeficiency Virus
HR	Human resources
IMF	International Monetary Fund
LA	Latin America
lndomex	Logarithm of the domestic public expenditures on HIV and AIDS per capita
MENA	Middle East & North Africa
MSM	Men who have sex with men
NASA	National Aids Spending Assessments
NGO	Non-governmental organization
NMAR	Not Missing at Random
NAA	National Aids Accounts
OCE	Oceania
ODA	Official Development Assistance
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
OOPE	Out-of-pocket expenditures
OVC	Orphans and vulnerable children
PHE	Public health expenditures
PLWH	People living with HIV and AIDS
PPP	Purchase power parities
PvtHE	Private Health Expenditure
SEA	South & South East Asia
SSA	Sub-Saharan Africa
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNGASS	United Nations General Assembly Special Session on HIV/AIDS
WCE	Western and Central Europe
WHO	World Health Organization

Contents

1. Introduction

Keeping track of the resource flow addressed to HIV (Human Immunodeficiency Virus) AIDS (acquired immunodeficiency syndrome), and the resources available is the first step when it comes to examining the effectiveness of the aid and deployment for an effective response. Until recently, estimates for HIV expenditures have only covered spending in the health sector (UNAIDS, 2006: 234) Expenditures on HIV and AIDS do not only consist of health expenditures, in fact it is estimated that more than half of the total amount required for the AIDS response each year should go to prevention (UNAIDS, 2006: 226). Expenditures on HIV and AIDS also include treatment and care, in which antiretroviral therapy (ART) requires the greatest share of money, support for orphans and vulnerable children, programme support and infrastructure, and human resources. Resource tracking is essential when determining the gap between resources available and resources needed. This analysis often reveals severe misallocations of the money. Making this visible is not only important when it comes to decision making, but also when it comes to policy dialogue and advocacy. For example, the Global Fund requires 10% to 20% of counterpart financing over the proposal term if the recipient is a lower-middle income country in round 7 (Global Fund, 2007)

1.1 Problem statement

Tracking flows allocated to HIV and AIDS is complex work as they come from various sources, the main ones being bilateral, multilateral, private and domestic sources. The focus of this project is to identify the level of domestic public expenditure, it is an important part of resource tracking; it currently consists of about 30% of the available resources for HIV and AIDS on a global level. Identifying the government's level of domestic spending gives an indication on its national commitment on HIV an AIDS. It is also essential when prioritizing aid recipients; when funds to HIV and AIDS are scarce they should be allocated to the countries needing them the most. All countries and donors should give a higher priority to reducing the burden placed on low-income countries. Knowing the level of spending of a country is the first step in examining the country's ability to take on a grater share of the financial burden to sustain a response to the HIV epidemic. Furthermore, donor flows are shown to be highly volatile and unpredictable (Lewis, 2005:12). Where it is possible, relying more on the countries' own resources and thus having a more predictable funding, will enable the government to have longer term projects.

One of the main problems in addressing the spending is the lack of information. Not all countries have a system to record their spending on HIV and AIDS, and if they do, it is not always obvious how to interpret the data. This project will on the one hand contribute to creating a database for government spending, and on the other hand to develop a model including economic, demographic and epidemiological factors among its explanatory variables. The model will identify the main drivers of the spending to make reliable estimates and predictions where data is not available. Previous work by the Resource Flows Project projected domestics government expenditure related to STD/HIV/AIDS largely based on GDP growth (Dalen and Reijer, 2006). Identifying the main explanatory factors for changes in domestic public expenditures will make the projection modelling more sophisticated and accurate, and provide predictions were no data is available. Currently we have information for 81 countries from 2004; the goal is to have predictions for 55 other countries. That is, the best data available for HIV expenditures from countries themselves will be used to make an econometric model to fill in the gaps for the 55 countries that do not have data available. Then, this total of 136 countries will give the estimated level of HIV expenditures in low- and middle-income countries.

1.2 Previous studies

According to UNAIDS, the Resource Flows Project was engaged to estimate and project resource flows from donor governments and governments and NGOs in developing countries and countries in transition. They have projected domestic public expenditures from 2005 to 2007 (Dalen and Reijer, 2006), with an estimated 3,062,675 USD spent by governments in 2005 in developing countries and countries in transition. They assumed that "projected funding levels grow according to the set longitudinal trend. Therefore, for projections in countries with no secondary information, the growth in funds in 2005 and 2006 is completely driven by developments in national income (GDP)." The Resource Flows Project used the same database as in this study (UNAIDS, 2006), except for the additional reports and update received after the publication of their paper.

2. Background information

2.1 Overview of the epidemic

HIV is a retrovirus that causes AIDS, which leads to failing immune systems making humans vulnerable to life-threatening opportunistic infections. Today, an estimated 33.6 million [33.4 million-46.0 million] people worldwide are affected with HIV according to the UNAIDS report on the global AIDS epidemic (UNAIDS, 2006:9). The nature of the epidemic is classified in three different categories: low level, concentrated or high level, emerging or advanced (UNAIDS, 2006:234). The transmission patterns vary among countries. Unlike Western Europe where most infections occur in men who have sex with men (MSM), heterosexual intercourse is the greatest trigger in Sub Saharan African countries. The three main transmission routes are unprotected sexual intercourse (as mentioned above), contaminated needles, and transmission from an infected mother to her baby at birth, or through breast milk (vertical transmission).

Joint United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organization (WHO) estimate that AIDS has killed more than 25 million people since it was first recognized on December 1, 1981, making it "one of the most destructive pandemics in recorded history" (UNAIDS, 2006:4). The fact that HIV is transmitted through body fluids makes it prone to stigmatization and challenges the response to the pandemic.

Countries need sustained and predictable funding that increases over time because of the long term nature of the epidemic. 5.5 billion people, or 85% of the world's population live in the 148 countries classified as low- and middle income countries in 2005, and more than 90% of people living with HIV are from these countries (UNAIDS, 2006). The estimated annual funding for the AIDS response in that year was USD 8.3 billions, compared to the estimated need of about USD 12 millions. That leaves a gap of about USD 4 millions that needs to be filled to reach Universal Access[1] to treatment for people who are affected.

[1] The requirements for reaching Universal Access have changed due to changes in the assumption of the net survival of people living with HIV without receiving treatment; this is 11 years instead of previously 9 years for all countries except those where subtype E constitutes the majority of infections. As part of this assumption,

Universal Access is not the same as 100% coverage in treatment, as there will always be some people not wishing to seek treatment of various reasons.

2.2 Resource mobilization

The best available estimates of global needs for AIDS indicate that global resource requirements amount to a total of USD 15 billion in 2006, USD 18 billion in 2007 and USD 22 billion in 2008 for prevention, treatment and care, support for orphans and vulnerable children (OVC), as well as programme and human resource costs (UNAIDS, 2005) (Table 1).

Table 1 AIDS Resource needs in US$ billions

US$ billion	**2006**	**2007**	**2008**	**Total for 2006–2008 ***
Prevention	8.4	10.0	11.4	29.8
Treatment and care	3.0	4.0	5.3	12.3
OVC	1.6	2.1	2.7	6.4
Programme costs	1.5	1.4	1.8	4.6
Human resources	0.4	0.6	0.9	1.9
Total	**14.9**	**18.1**	**22.1**	**55.1**

* The totals for 2006–2008 have been rounded to the first decimal place, with the result that there may be small differences with the figures for subtotals in the text because of rounding errors.

The estimates for the AIDS resource needs are based on the High-Level Meeting on "The Global Response to AIDS: Making the Money Work – The Three Ones in Action" on March 9th 2005 (UNAIDS, 2005). The figures have then been refined and updated on the basis of newly available information and input from the recently formed Resource Needs Steering Committee and the Resource Needs Technical Working Group[2]. Currently, the Resource Needs Technical Working Group is working on updated estimates, including new interventions like male circumcision. These are scheduled to be out by June 2007.

people living with HIV will become eligible for treatment an average of 3 years before they are expected to die from an AIDS related cause, instead of previously 2 years (Stover, 2007).

[2] These advisory groups were formed subsequent to March 9 2005, and are composed of international economists and AIDS experts from donor and developing countries, civil society, United Nations agencies and other international organizations.

Lately, there have been some concerns in the media that too much money has been spent on HIV and AIDS (England, 2007). Paul Delay and colleges argue that in fact the USD 15 billions needed in 2006, only USD 9 billions were available for HIV and AIDS (Delay et al., 2007). Furthermore, they claim that "the bulk of the funding is additional to amounts spent on other aspects of health development", reflecting that money for HIV and AIDS are used to build up health infrastructure. They also stress the importance of that "greater efforts are needed to make sure that those countries that are able to do so invest more of their own money in AIDS".

2.3 Governmental effects

Receiving substantial amounts of HIV/AIDS funding can also affect the macroeconomic stability of a country (Lewis, 2005). A large aid flow could potentially lead to appreciation of the exchange rate and an increased inflation rate. The absorptive capacity is generally low in low-income countries, and there is a problem that resources are diverted from other sectors (potential Dutch disease). A problem with receiving aid is that the government often does not have the saving/investment decision, since most aid comes with time limits and restriction on its use. These are negative short term effects on the economy. In the long term, macroeconomic literature suggests that if the aid is used effectively this effect can be reversed. Prolonging the life of the citizens increases the country's human capital and hence its production capacity. However, aid flows are shown to be highly volatile (Bulir & Hamann, 2003), hence, dependency on aid flows is risky for the sustainability of the response to the HIV epidemic. If the governments are capable of taking on more of the responsibility themselves instead of relying on aid the planning of long term objectives will become easier, and thus problems of unpredictability and inconstancy often incurred with receiving aid can be avoided.

3. Data sources, data, and definitions

In this section, the variables included in the analysis will be defined, as well as the conceptual framework giving the reasoning of the analysis performed in section 4.

3.1 The dependent variable – domestic public expenditures on HIV and AIDS

The definition of the variable

The dependent variable, the domestic public expenditure related to HIV and AIDS, is defined as the amount of national funds (including reimbursable loans) spent by governments from domestic sources for HIV and AIDS. This includes expenditures on a regional and general level, as well as social schemes. The expenditures on HIV and AIDS do not include only health related issues, but also other issues as programme costs and advocacy in schools and workplaces. Significant spending has also been directed to strengthening of health infrastructure and capital, e.g. laboratory networks, universal precautions, blood bank safety and safe injections.

The main spending categories are:

1. Prevention
2. Care and treatment
3. Orphans and vulnerable children[3] (OVC)
4. Programme management and administration strengthening
5. Incentives for human resources
6. Social protection and social services (excluding OVC)
7. Enabling environment and community development
8. Research (excluding operations research included under programme management)

Sources of information

The primary database for the domestic public spending would be from the UNGASS country reports which 189 countries have committed themselves to deliver every two year. It addresses global, regional and country-level responses to prevent new HIV infections,

[3]In the context of resource needs estimates and AIDS Spending Assessments, vulnerable children are defined as those who have at least one parent who is alive but seriously ill (mainly because of HIV) and unable to take care of them.

expand health-care access and mitigate the epidemic's impact. At the United Nations General Assembly Special Session on HIV/AIDS (UNGASS) in June 2001, the Declaration of Commitment on HIV/AIDS was adopted by 189 Member States (UNAIDS, 2007). The Declaration of Commitment reflects global consensus on a comprehensive framework to achieve the Millennium Development Goal of halting and beginning to reverse the HIV/AIDS epidemic by 2015. The information is summarized in the UNGASS 2005 Reporting (UNGASS, 2006), and the domestic public expenditures used to be reflected in first indicator (GE-01) among a series of core indicators developed to measure the progress in implementing the agreed-on commitments from the Declaration. After the revision on the indicators issued March 2007 (UNAIDS, 2007), this indicator will now consist of domestic expenditures by categories and financing source. The data collection frequency is annual and the National Aids Spending Assessment will be the primary tool of measurement.

When the UNAIDS Progress Report on the Global Response to the HIV/AIDS Epidemic (UNAIDS, 2006) was published, all of the country progress reports were not yet returned and new arrivals will be used to fill in numbers if they are sufficiently explained. The country progress reports can also be handed in using the Country Response Information System (CRIS), which is a computerized and standardized complement to the UNGASS narrative report. Another source is the National Aids Spending Assessments (NASA). This is a fairly new initiative and not many reports are received yet. NASA is a country-based survey sent to governments, reporting on HIV/AIDS spending, which is not only related to the health sector. The NASA predecessor, National AIDS Account (NAA) is an important source for earlier years. The NAA tracks the total HIV/AIDS expenditures from all sources of finance, including public, private, households and external sources. This data is in the SIDALAC database which mostly includes Latin American countries. The fifth source for data is from the National Health Accounts which has a sub analysis for HIV/AIDS. The last source would be various financial economic reports with relevant data, to be specified later in the paper. The database for this work was retrieved using all the six sources of information.

Limitations

There are limitations to the data available. Good and accurate data has been available only in the last years. Reporting takes time and capacity, which the governments, especially in low-income countries, lack. When countries reported public expenditures, it has not always been clear what is meant. Instead of the actual amount spent, some reported commitments and some reported the amount budgeted, which can differ severely. In some cases, there was also a problem of what the number contained; public expenditures related to HIV/AIDS mixed together with aid, or they only included the HIV expenditures related to health, etc. This could be due to the survey form, or perhaps limited time to read the manual. However, the data chosen for this analysis was thoroughly screened, and many reports were rejected. All the data for this analysis was validated by the countries themselves. Many countries have not reported information on HIV expenditures at all, for the UNAIDS Progress Report on the Global Response to the HIV/AIDS Epidemic (UNAIDS, 2006). Only 40% of the countries that submitted reports supplied information related to national indicators, such as the GE-01 indicator.

Populating the dependent variable

With 65 data points, 2004 is the most complete year, followed by 2003 and 2005. Thus, 2004 was used for the further analysis. However, in some cases the adjacent years were used to populate horizontally using averages or linear interpolation and extrapolation according to certain criteria.

The data was populated horizontally according to some conservative rules: only missing data for the year 2004 were filled in to increase the sample size for the cross sectional analysis. The missing data between years where observations are available were filled in; If there at least as many observations available as the number of missing data, the blanks were replaced by estimates using linear interpolation (n=4). When there were only two observations the average was used to fill in the gap (n=2). Extrapolations ahead were only done when having four observations or more. If there were observations for four years only one year ahead will be estimated (n=3). Two years ahead were estimated when having five or more consecutive years (n=2). There were no estimations more than two years ahead. The method used was linear extrapolation (least squares). Two missing data were replaced by average because the trend component was too large to use linear extrapolation, and three were replaced by the average of the observations for 2005 and 2006.

In total, the year 2004 had 65 data points for public domestic expenditures originally, and after the estimation methods described above the sample size increased to 81 data points, and hence making it more robust.

3.2 The conceptual framework: theory and explanatory variables

The aim of the empirical analysis is to explain what determines the domestic public expenditure on HIV and AIDS for a country. Another goal is to predict these expenditures in countries where data for this variable is lacking. Considering that expenditures on HIV and AIDS include prevention, treatment and care, different explanatory variables were included. Treatment is likely to be driven by the burden of the disease, while prevention expenditures would be depending on the HIV awareness in the country. In both cases, the budget constraint is limiting the spending possibilities, as well as other constraints the countries face.

Through discussions with experts in UNAIDS, four different variables were identified and thought to be collectively exhaustive in describing what determines the HIV expenditures in a country:

1. *The amount of money available*: The government expenditures on HIV and AIDS most likely depend strongly on their income. The income affects the level of spending because, obviously, the governments cannot spend more than they have, and often the income is not sufficient to cover the need in low- and middle income countries.
2. *The ability to spend the money available*: Evidence from Sub-Saharan Africa suggests that even though they have sufficient funding, Sub-Saharan countries have difficulties in using the money available because of constraints on the absorptive capacity. One constraint could be that they do not have enough human resources (HR), another would be the lacking of infrastructure. Another measure that affects the ability to spend available money is governance and stability. The economic environment is crucial when it comes to ability of spending, as a country in conflict or post conflict situation, or with corruption problems is not as capable of spending the money to meet public needs like other countries with a different political reality.
3. *The burden of the disease*: If a government allocates money to HIV and AIDS it is because people in the country are affected by HIV and AIDS. Countries with a high burden of the disease need more resources than countries with fewer HIV cases, and thus are prone to allocate more resources to HIV and AIDS. However, countries most strongly affected by HIV often are the countries with the least resources.

4. *The willingness to spend the money available*: Some governments might have the money to spend and the ability to spend the money on HIV and AIDS, but lack the political will or commitment. In Latin America there is a saying that "demand drives the expenditures on care and treatment, but political will drives prevention expenditures".

These four variables that, according to this simple theory, would explain HIV related expenditures are difficult, if not impossible to observe as such. We would therefore have to seek for observable variables carrying information on these four unobservable variables.

Figure 1 The conceptual framework

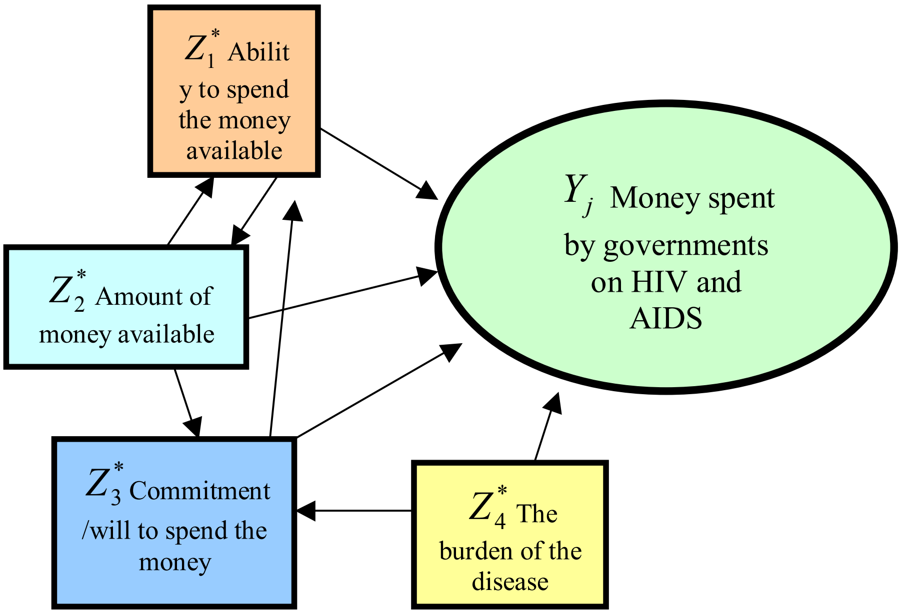


Figure 1 describes the relationship we assume between the four unobservable variables, to be considered as exogenous, and the HIV expenditures. The arrows indicate our belief that there is a uni-directional causality in that expenditures on prevention and treatment have no feedback on the "need for the money". Neither do we intend to model that these expenditures in the current year may affect the burden of disease in later years. We make this simplification because we have one observation from each country only, and hence are forced to use a static model. The arrows between the four exogenous variables indicate that they are not unaffected by each other. The amount of money available is expected to affect the ability to spend money, since richer countries often have more human resources or less problems with for example corruption. The same goes for the inverse relationship; the better infrastructure and less corruption, the better environment for generating money. The amount of money available is also expected to affect willingness to spend money, since it is easier to commit to spending money, or express political will to spend money, if there is actual money to be spent. Willingness to spend money is expected to have an effect on the ability to spend money, since a committed government would build up the infrastructure needed to enable spending. Last, the arrow between the need and willingness to spend money suggests that the stronger the burden of disease, the more pressure on the government to pledge for spending money on HIV interventions. These relationships between the four exogenous variables are not, however, part of our formalized model.

The four latent variables were assumed to have a linear relationship with the domestic public expenditures on HIV and AIDS. The relationship can be described by the following equation:

(i) $Y_j = a_0 + a_1 Z^*_{1j} + a_2 Z^*_{2j} + a_3 Z^*_{3j} + a_4 Z^*_{4j} + u_j$

Where Y is the domestic public expenditures on HIV and AIDS for country j and Z^*_{ij} (I = 1,2,3,4) are the four variables. The a's are constants, which would be bigger the more important the variable is. As explained above, our theory suggests that the variables all have a positive relationship with the HIV expenditures, and thus conceptually the a's are all positive. u_j is the normal distributed error term, capturing all the factors that cause domestic public expenditures to vary not captured elsewhere in the model. This equation rests on the assumption that the Z^*_{ij} are exogenous, that there is no feedback from the Y_j to Z^*_{ij} (indicated by the arrows in figure 1 only pointing towards the money spent by governments on HIV and AIDS, and not backwards), which gives $\operatorname{cov}(Z^*_{ij}, u_j) = 0$ for all i and j.

The problem is, however, that since the Z^*_{ij} are non-observable, or latent variables, the a's remain unidentified and the domestic public expenditures on HIV and AIDS cannot be predicted from equation (i). Latent variables, as opposed to observable variables, are variables that are not directly observed but are rather inferred from other variables that are observed and directly measured, called proxies. For a variable to be a good proxy it must have a close connection (or high correlation) with the inferred value. Thus, good proxies for $Z_{1j}^*,...,Z_{4j}^*$ must be found for as many countries j as possible. Assuming that the proxies found only affect the according latent variable, thus that the latent variables are mutually exclusive, the relationships between the proxies and the latent primary explanatory variables could be modeled by the equations:

(ii) $Z_{ij}^* = \alpha_0 + \sum_{k=1}^{k_i} \alpha_{ik} X_{kij} + e_{ij}$,

where i = 1, 2, 3, 4, j is still the country number, and k indicates the number of the proxy X, aiming to explain the latent variable Z_{ij}^*. α_{ik} cannot be known or estimated since Z_{ij}^* is unobservable. e_{ij} is the error term capturing effects of changes in Z_{ij}^* not explained by the X's.

Combining (i) and (ii) to eliminate $Z_{1j}^*,...,Z_{4j}^*$, gives an equation between observable variables:

(iii) $$Y_j = \underbrace{a_0 + a_1\alpha_{10} + ... + a_4\alpha_{40}}_{\text{int}\ ercept} + \underbrace{a_1 \sum_{k=1}^{k_1}}_{estimated} \alpha_{1k} X_{k1j} + ... + \underbrace{a_4 \sum_{k=1}^{k_4}}_{estimated} \alpha_{4k} X_{k4j}$$

$$\underbrace{+u_j + a_1e_{1j} + ... + a_4e_{4j}}_{error=v_j},$$

where the coefficients $a_i\alpha_{ik_1},...,a_i\alpha_{ik4}$ are the coefficients to be estimated in the multiple regression model. The intercept and the error term are now combinations of the intercepts and error terms in equations (i) and (ii). The number of observable regressors will be *k1+k2+k3+k4*.

Some other assumptions:

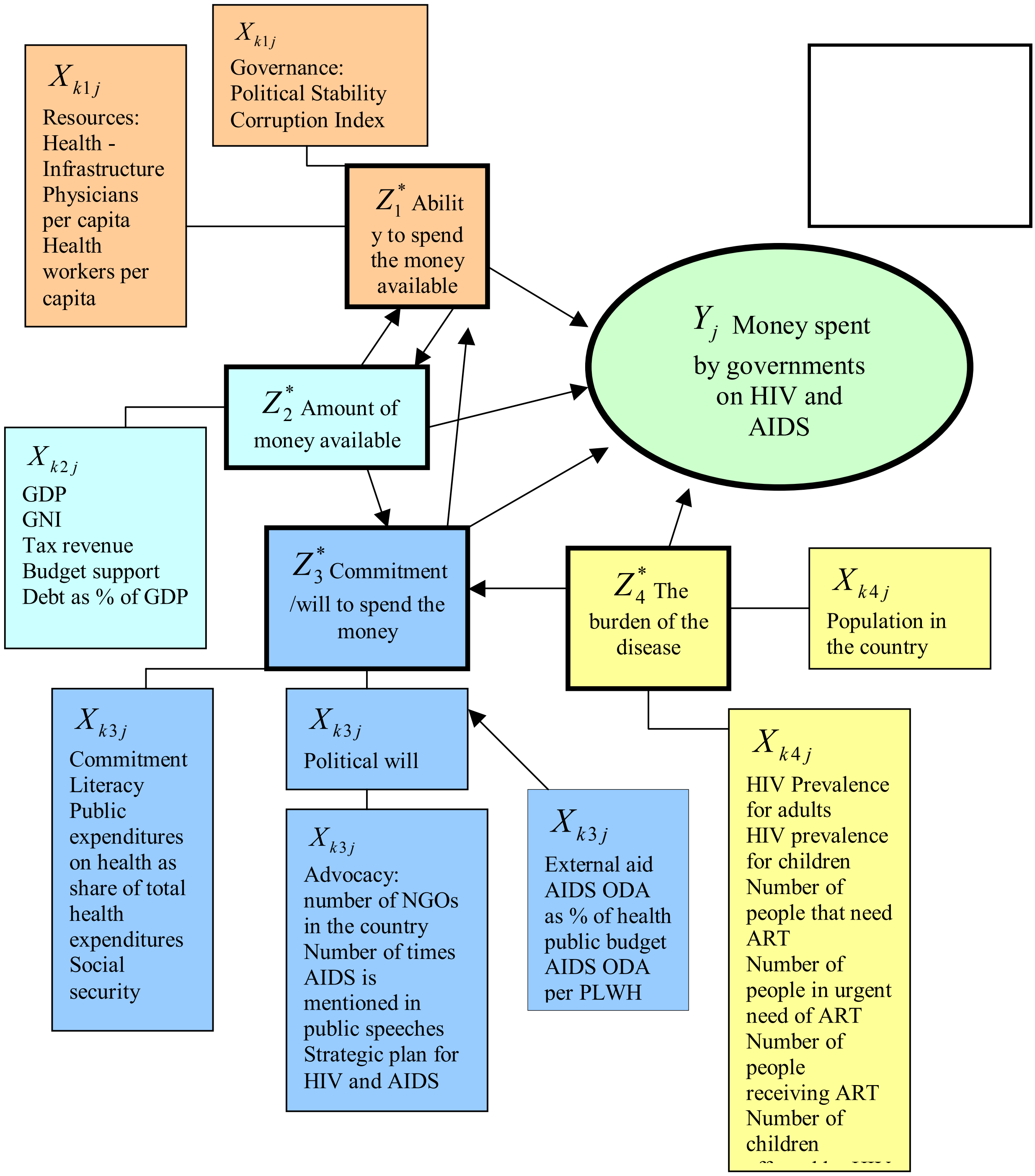
1. u_j is uncorrelated with $(e_{1j},...,e_{4j})$
2. u_j and $(e_{1j},...,e_{4j})$ are uncorrelated with all X_{kij}

These assumptions make the new error term v_j uncorrelated with the all X_{kij}, so again no feedback is assumed from the Y's to the X's.

Selection of proxies – further discussion

Figure 2 visualizes the extended conceptual framework when all the proxies (X's) are included. The availability of the different proxy variables, their possible overlapping, and how to proceed with distinguishing between the indispensable proxies and those which can be represented by others, will be discussed during the empirical analysis in order to find the variables that should be tested when searching for the final model.

Figure 2 The conceptual framework with proxy variables indicated



In principle, all Z's could depend on all X's, but in this model the proxies are assumed to affect only one X each. The exception is the share of people in urgent need of ART that receive ART treatment, which could carry information on the ability to spend money and the willingness to spend money.

The explanation of the choice of proxies is divided in sections of the four latent variables they are aiming to explain:

The burden of the disease: A proxy for the burden of the disease in the country should be included in the model. Information suggests that when prioritizing government funds, children have the highest priority, followed by women, and last men. Public funds are more easily raised to save children, which would suggest that prevalence among youth could be a determinant for the public expenditure on HIV and AIDS. However, governments could be reluctant to spend money on HIV and AIDS on adults, but when the urgent need is great they face no other choice. When a large ratio of the working people is affected by HIV, spending money on HIV has higher priority so people in urgent need of ART treatment should be tested as a determinant for HIV expenditures.

Amount of money available for spending: The gross domestic product (GDP) or gross national income (GNI) of a country is often used to indicate a country's income; it is easily obtained and measures how much the economy actually produces within a country. In an article by Musgrove, Zeramdini and Carrin (2002), the GDP level has shown to be associated with total health spending. The analysis included both high and low income countries, and one result was that total health spending rouse from around 2–3% of GDP at low incomes to 8–9% at high incomes (Musgrove et al., 2002:134), and that health takes an increasing share of total public expenditure as income rises. Since a part of the public HIV spending is included in the public health spending, GDP is suspected to be correlated with domestic public expenditures on HIV and AIDS. The choice has to be made whether the GDP should be calculated using purchase power parities (PPP), international dollars or average exchange rate. Converting to PPP would give a better picture of the country's purchase power of the available HIV and AIDS expenditures. For instance, an equal amount of money would pay for more physicians in one country than in another. If, however, the GDP is converted, the domestic public expenditures should be converted in the same way.

When it comes to representing the money a government has available for spending, GDP alone might not be the solution. A country with a large budget deficit can at the same time have a high GDP. Taxation is one source of income for the governments, and the percentage of government income which comes from tax could improve the measurement of their actual income. However, since tax is a less important source of revenue in low- and middle income countries than in high-income countries, perhaps it is not that relevant in our model.

The amount of aid given as budget support would increase the money available in the country. Budget support was chosen to capture the effect of aid flows increasing the spending capacity of the government. While aid received by the country is often earmarked certain projects, budget support is given directly to governments and is undistinguishable from other sources of revenue and should therefore be included in the analysis.

The external debt service information we tried to track down was on how much of the money available for the government went to debt. Many countries claim having problems since a substantial part of the revenues goes to paying debt, and the variable should therefore be relative to the income of the country. The variable found was debt service as percentage of GNI. The reason for using debt service was that poor countries had a large percentage of total debt of the GNI, but smaller debt service. That could be interpreted as poor countries with a large debt, but not paying off a large amount each year. For this analysis, it is better to use the percentage they actually pay, since the goal is to examine how the money available each year is affected by having to pay debt. A country using 70 percent of the GNI to pay back debt would be assumed to have less money to spend on a response to HIV.

Ability to spend money: Since a country's absorptive capacity is difficult to measure, and currently there are no existing indexes available, variables carrying information on the level of absorptive capacity in the country should be included in the model to account for this effect. The number of physicians per capita gives an image of the human resources in the country. Since it is not necessary to be a trained physician to give antiretroviral treatment or to perform other services needed, the number of health workers or nurses per capita could be a better measurement.

Governance could be another issue impeding the ability to spend money. If the country is politically unstable it is difficult to deliver services, and infrastructure may break down. A WHO report on the African region for instance states: "Social and political instability, including war and mass population displacement in several countries, have disrupted health services" (WHO, 2002:65). The more politically unstable a country is, the more unable the country is to spend money on HIV and AIDS; one reason being that unstable countries in conflict have high military expenditures. This indicator is suspected to be correlated with GDP, since poor countries are often more politically instable. There is also evidence that HIV prevalence is associated with poor governance (Menon-Johansson, 2005).

Another variable that affects the ability to spend available money is the level of corruption. Money intended to go to HIV and AIDS projects could go elsewhere due to corruption. Tayler and Dickinson (2005:10) identify treatment programmes as the most vulnerable, since "Money for high-value drugs can be embezzled at any number of points in the procurement and distribution chain." This could give an indication on how the government policy acts in the country and it also provides a hint of the absorptive capacity.

Commitment and willingness to spend money: The level of commitment of a country aims to measure how much a government is concerned about its citizens; if it usually spends a lot on social security relative to other expenditures, it could be more prone to spend money on HIV and AIDS. However, no measure of social security per se was found. Since education is a large share of social spending, the literacy rate was used as a proxy for how much governments spend on education, since a high literacy rate would imply a good education system and thus a higher level of spending on education.

The size of the public health expenditures is also an indicator of how much the government commits to paying for its citizens. A government willing to allocate money to the health sector would be expected to be more prone to spending money on HIV and AIDS than governments that do not. However, the care and treatment part of HIV and AIDS spending would be included in the public health expenditures. Since antiretroviral treatment is expensive, this is often a great share of the total spending for HIV and AIDS. One would also believe that total health expenditures are strongly correlated with GDP per capita.

The share of out-of-pocket expenditures (OOPE) on health could give indicate how committed the government is. Higher public expenditures have been identified as important financing mechanism to bring down the share of out-of-pocket expenditures (Xu et al. 2003), so a higher share of out-of-pocket expenditures should mean lower public domestic expenditures. An inverse of the OOPE share could then be a predictor for how much the government spend in a country on HIV and AIDS. The OOPE for health is not the same as for HIV and AIDS, but currently good estimates are not available for HIV and AIDS. Most likely, OOPE for HIV and AIDS are higher than for health in general because catastrophic diseases are usually less covered by insurance, particularly by private insurance. Since the estimations for OOPE on health are survey based and most likely not as accurate for all countries, using the domestic public health expenditures as share of total health expenditures is better. Instead of using domestic public health expenditures per capita in the model, the share of the total is used. Using the share of the domestic public spending and not the amount the model captures whether the burden on people is relatively high or low, and thus avoiding the use of OOPE share.

However, countries could spend a lot on education and other social services, but they could be reluctant to spend money on HIV and AIDS. The reasons could be stigma or ignorance about the epidemic, or perhaps that HIV and AIDS is not prioritized. Thus, the level of advocacy in the country could serve as an indicator of how much awareness they have. In this respect, it might be useful to look at how many NGOs there are in the country. NGOs do lobbying to make governments increase funding for, and prioritize HIV and AIDS. They also raise awareness about HIV and AIDS, escalating pressure from the people to act on giving a response to the epidemic.

The government spending on HIV and AIDS is also affected by how much external aid the country receives. External aid could be a disincentive in the way that the country use less money from the government budget, since funding is received from elsewhere. When governments know funding would be high for HIV and AIDS, they are likely to prioritize other matters. Official Development Assistance (ODA) could be measured as a share of public health expenditures, because if some countries receive as much ODA for AIDS as their total health budget, they might think that AIDS should not be prioritized, but rather divert funding towards other issues.

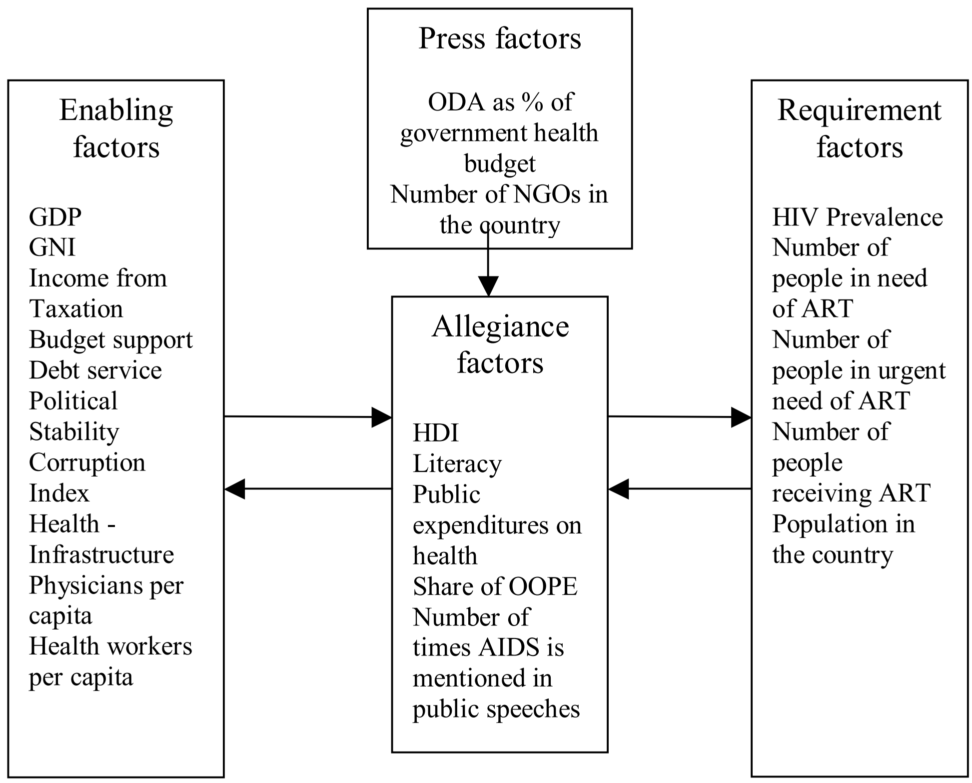
However, International funding channels like the Global Fund to Fight AIDS Tuberculosis and Malaria (GFATM) emphasize that their funds should be complementary and additive in the way that they should not replace existing national and international resources (Global Fund, 2003). The Global Fund also encourages political commitment concepts, like per capita health expenditure, and existence of supportive national policies and multiyear strategic plans. Hence, countries might make an effort to show political will and government spending to attract donors. Moreover, a high level of external aid could also mean that the awareness of HIV is elevated, and that HIV spending is put in focus and prioritized. This would suggest that the negative effect could be minor, or in fact this effect might even be positive.

In conclusion, a priori, it is difficult to say if the effect of external aid will be positive or negative, but one would think that the first argument has a greater impact on the decision makers; i.e. the government, resulting in that the expected effect of high external aid would be negative.

The internal dynamics

Figure 3 below captures the internal dynamics in the model as discussed briefly above. The arrows represent supposed time lags which are not included in our static model. *Enabling factors* includes the economic, demographic and other macro development indicators giving information on how able the country is in spending money on HIV and AIDS. *Allegiance factors* give an indication of how much the country invests in its citizens, or how devoted it is in fighting HIV. The arrows between them go both ways: the country supports its citizens more when the government has more resources, and the more the country has an obligation to its citizens, the more they spend of their resources. *Requirement factors* are the epidemiological ones indicating the actual need in the country for HIV and AIDS spending. The arrows are going both ways in this case as well. The need is affected by how large resources the country has and through the allegiance factors how much it spends, because the more it spends, the more people are treated and the more people are prevented from getting HIV. The effect is also reversed since countries often spend more on HIV and AIDS if the burden of disease is high (and they are able to). Finally, *press factors* are indicators that could affect the allegiance or commitment the country has for the people affected by HIV. The NGOs in the country are raising awareness about HIV and AIDS, creating an incentive for the government to spend more money on HIV and AIDS. External aid, however, could be a disincentive to use money, since the government is not the sole provider of the money and they could delegate the money to other budget posts. The aid effect can be both negative and positive because of the previously discussed additive principle of the Global Fund. Thus one cannot be completely sure of the sign of this parameter.

Figure 3 Internal dynamics between predictors



3.3 The definition of the proxy variables

The explanatory variables stated in the conceptual framework were not all available, and for those that were, the reliability of the information was considered when choosing explanatory variables. The variables were not adjusted for purchase power parities (PPP) because the domestic public expenditures are not adjusted for PPP's. This conversion should be made in order to compare the countries' expenditures, but for purpose of this analysis the data was not converted.

All variables are for the year 2004 if not stated otherwise. This choice was motivated partly by how they are thought to affect decisions, and partly by their availability. When the data was scarce, the latest year available was used. Table A1 in the Annex includes an overview of how observations of the explanatory variables were retrieved. For example, the Tax Revenue variable had 50 data points for 2004, but 76 data points when including the latest year available. Whenever an observation was reported as a range, the midpoint was used (for example >0.1 was replaced by 0.05).

The proxies for Z_1^* : the ability to spend the money available

Health infrastructure

- Number of hospital beds (per 10 000 inhabitants) was retrieved from the WHO database of Core Health Indicators. Hospital beds include in-patient and maternity beds. Maternity beds are included while cots and delivery beds are excluded. 41 data points were available for 2004, and 91 when the latest years available were used.
- Nurses (density per 1 000 inhabitants) was retrieved from WHO database of Core Health Indicators. Nurses include professional nurses, auxiliary nurses, enrolled nurses and other nurses, such as dental nurses and primary care nurses.54 data points were available for 2004, and 136 when the latest years available were used.
- Physicians (density per 1 000 inhabitants) was retrieved from WHO database of Core Health Indicators. Physicians include generalists and specialists. 52 data points were available for 2004 and 137 when the latest years available were used.

Governance

- Political stability is one of the World Bank's six dimensions of governance, the others being voice and accountability; government effectiveness; regulatory quality; rule of law; and control of corruption. Initiated in the late 1990s, its six dimensions are based on a longstanding research program of the World Bank Institute and the Research Department of the World Bank (World Bank, 2006). The aggregate indicator is constructed using an unobserved components methodology. The indicators are measured on a scale ranging from about -2.5 to 2.5, with higher values corresponding to better governance outcomes. The political stability was chosen since the dimensions of governance are highly collinear; around 0.8 in this sample, and since political stability was the one most correlated with the dependent variable (table A2). In the final model all the World Bank's variables of governance will be tested for significance. The political stability indicator reflects the statistical compilation of responses on the quality of governance (How does risk of political violence influence government?). It gives the perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism (Kaufmann et al., 2003). The indicator is based on 31 data sources produced by 25 different organizations worldwide, e.g. African and Asian Development Bank, the Economist Intelligent Unit, Business Environment Risk Intelligence and the World

Economic Forum. The indicator contained negative values, thus the range of the distribution was moved from (-3.055, 1.448) to (0, 4.503) to ease the interpretation of the regression.

- The Corruption Perception Index – The Corruption Perception Index by Transparency International (Transparency International, 2004) gives a measurement of the degree to which corruption is perceived to exist among public officials and politicians. The TI Corruption Perceptions Index (CPI) ranks more than 150 countries by their perceived levels of corruption, as determined by expert assessments and opinion surveys. The index might come under allegiance as being very subjective, as it is survey based and it reflects the views of businesspeople and analysts from around the world, including experts who are resident in the countries evaluated, but it is the most suited measure of corruption that can be used for empirical studies.

The proxies for Z_2^* *: the amount of money available*

- Gross Domestic Product (GDP) was obtained from the World Development Indicators (World Bank database). It measures how much an economy actually produces within a country. The limitations of to this measure are that developing countries often have a large informal sector, with transactions not being registered directly as output in official statistics. World Bank, data on population for each country was used to convert GDP and the domestic public expenditure on HIV and AIDS to per capita values.
- Gross National Income (GNI) was retrieved from the World Bank's World Development Indicator database. GNI measures the total value of goods and services produced within a country together with its net receipts for primary income received from other countries. GNI measures the countries claims on output rather than actual output as GDP does.
- Definition: GNI (formerly GNP) is the sum of value added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad. Data are in current U.S. dollars. Source: World Bank national accounts data, and OECD National Accounts data files. Data from the World Bank on population was used to convert the figures to per capita values.
- Debt service as percentage of GNI: The variable was retrieved online from the World Development Indicators database. Total debt service is defined as the sum of principal repayments and interest actually paid in foreign currency, goods, or services on long-

term debt, interest paid on short-term debt and repayments (repurchases and charges) to the IMF. The GNI denominator is a three-year average.

- Tax revenue as percentage of GDP was retrieved from the World Bank's World Development Indicators database. According to the information reported, tax revenue refers to compulsory transfers to the central government for public purposes. Certain compulsory transfers such as fines, penalties, and most social security contributions are excluded. Refunds and corrections of erroneously collected tax revenue are treated as negative revenue.
- Budget support in 2004 (2004 in current US dollars) was retrieved from the Organization for Economic Co-operation and Development (OECD), Creditor Reporting System (CRS). Budget support is money given as Official Development Assistance, none earmarked directly to the national treasury, hence undistinguishable from other country revenue.

The proxies for Z_3^**: the willingness to spend the money available*

Commitment

- The variable on total domestic expenditures on health is a per capita measurement on domestic health expenditures for low- and middle-income countries, taken from the WHO database of Core Health Indicators. They are not necessarily the official statistics of Member States, which may use alternative methods. To date, only data from 2003 is available, thus the health expenditures converted by an average exchange rate was used to make them more comparable with the expenditure on HIV and AIDS.
- Literacy rate, as it is measured by the UN, gives the percentage of the population between 15 and 49 years old which are able to read.
- Government expenditures on health at average exchange rate (US$) as % of total health expenditures is measured as general government expenditure on health as percentage of total expenditure on health was retrieved from the WHO database of Core Health Indicators. The definition of general government expenditures on health (Public Health Expenditure) is the sum of outlays on health paid for by taxes, social security contributions and external resources (without double-counting the government transfers to social security and extra-budgetary funds). Total health expenditure is by definition the sum of Public Health Expenditure (PHE) and Private Health Expenditure (PvtHE). In

one case, the reported figure was below 1 for both government expenditures on health per capita and total health expenditures per capita; the midpoint 0.5 was then used.

- Official Development Assistance (ODA) Commitments in 2004 (2004 in constant dollars) divided by domestic public health expenditures. The ODA statistics were retrieved from the Organization for Economic Cooperation and Development (OECD), Creditor Reporting System (CRS). Commitments were chosen over disbursements since this is known by the government at the beginning of the year, when they budget how much they should spend on HIV and AIDS, whereas the actual disbursed amount is not known until the end of the year. When no figure was reported, zero was used (n=35), assuming the country did not receive ODA if no commitments were reported. Three data points were lost due to lack of information on health expenditures. To include another variation of this variable, the ODA Commitments in 2004 (2004 in constant USD) was also divided by the number of people living with HIV.
- People with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%) was retrieved from the WHO database of Core Health Indicators. The definition is 'the percentage of people with advanced HIV infection receiving ARV therapy according to nationally approved treatment protocol (or WHO/Joint UN Programme on HIV and AIDS standards) among the estimated number of people with advanced HIV infection'. When the number was stated as <50, the rule of using the midpoint, 25, was applied (ten cases). This data was recently published and is only available for the year 2005, but there is no reason to believe that drastic changes happened in one year. However, the endogeneity problem might be greater considering the year of the HIV expenditures is 2004, and the data on coverage for ART are 2005.

Advocacy

None of the variables for advocacy were found for 2004, and the decision was made not to use observations from later years since changes happen fast in this area.

The proxies for Z_4^* *:the burden of the disease*

Retrieved from the UNAIDS internal database, the following proxies for the burden of the disease are considered:

- Estimated number of people living with HIV, children (0-14)
- Estimated number of people living with HIV, adult and children
- Estimated number of people living with HIV, adults (15+)
- Estimated HIV prevalence, children (0-14)

- Number of people with advanced HIV infection (urgent need)
- HIV prevalence - The prevalence is a UNAIDS measurement giving the percentage of the adult population (15-49 years) having the HIV virus. For Serbia, data on the prevalence for 2004 was lacking and the average prevalence between 2003 and 2005 was applied.

Table 2 present the summary statistics of all the proxy variables.

Table 2 descriptive statistics for the proxy variables

					----------	Quantiles	----------	
	n	Mean	Std.Dev.	Min	0.25	Median	0.75	Max
GDP per capita	131	2553	2682	91	481	1294	3794	10794
Budget support per capita	138	6.9	48.5	0.0	0.0	0.0	2.4	568.1
Total debt service (% of GNI)	124	62.5	35.9	1.0	31.5	62.5	93.5	124.0
Tax revenue (% of GDP)	76	16.6	8.3	2.0	11.4	15.3	21.0	54.4
Number of hospital beds (per 10 000 population)	93	34.2	27.8	1.5	12.0	25.0	53.0	132.0
Nurses (density per 1 000 population)	136	2.5	2.5	0.1	0.6	1.6	3.8	11.6
Physicians (density per 1 000 population)	137	1.2	1.3	0.0	0.2	0.6	1.9	5.9
Literacy rate	138	20.1	23.9	1.0	1.0	4.5	39.0	81.2
Political stability	137	-0.4	0.9	-3.1	-1.0	-0.3	0.2	1.5
Corruption Index	108	3.2	1.2	1.5	2.3	2.8	3.8	7.4
Government expenditure on health at average exchange rate (US$) as % of total health expenditures	136	0.5	0.2	0.2	0.4	0.5	0.7	0.9
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of domestic public health expenditures	135	0.2	0.4	0.0	0.0	0.0	0.1	2.7
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of people living with HIV	120	753	3594	0.0	1.5	37.1	170	29744
People with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%)	122	28.2	25.3	0.0	7.0	24.5	39.0	100.0
Estimated number of people living with HIV, children (0-14)	115	18841	43147	0.0	63	1374	13720	220000
Estimated HIV prevalence, children (0-14)	113	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Estimated number of people living with HIV, adult and children	127	280000	770000	0.0	4404	30932	170000	5500000
Estimated number of people living with HIV, adults (15+)	126	260000	740000	0.0	4383	31284	160000	5400000

HIV prevalence, adult (14-49)	127	2.7	5.6	0.0	0.1	0.7	2.3	33.1
number of people in need of ART	113	55450	140000	1.0	1113	8624	37352	960000

3.4 Design

The analysis is formally a cross sectional analysis for the year 2004 with 81 low- and middle income countries (ranked according to the World Bank Atlas method[4]) for which information on domestic public expenditures is recorded. Upper middle-income countries are also included; this increases the sample size and ensures more accuracy in estimation by giving a wider spread of the data.

A cross-section analysis will be used because it is difficult to do a panel data analysis since the range in years is limited. We only have reliable information for the past few years because of the scarce availability of data, due to lack of registration of strategic information. This implies that time-serial information will not be captured by the data set. For the present, we do not have enough data available for conducting a reliable panel data study.

Various variables will be tested to identify drivers for HIV and AIDS spending. An Ordinary Least Squares multiple regression model will be used to identify the final model from which we can predict domestic public expenditures related to HIV and AIDS for any given country.

[4]Economies are divided into groups according to 2005 GNI per capita, calculated using the World Bank Atlas method. The groups are: low income, $875 or less; lower middle income, $876 - $3,465; upper middle income, $3,466 -$10,725; and high income, $10,726 or more.

4. Analysis

The possible explanatory variables were collected, each considered a candidate for being an explanatory variable, and the dependent variable was analyzed to see their distribution, check for outliers and data entry errors, and look at the association of each variable with the domestic public expenditures for HIV and AIDS. Necessary transformations and tests were conducted to perform the analysis.

A bivariate analysis was performed initially to determine which explanatory variables were to be included as candidates for the multiple regression model. The variables were considered candidates if the p-value of the simple regression was 20% or below, or if the variable was considered important enough from the conceptual framework perspective. These variables will be included in the model building, and categorical variables will be tested, such as ‘regions’, since we expect different regions to have a different spending pattern.

When the final model was identified, diagnostics were computed to check the underlying assumptions. Unusual and influential observations were identified, and the assumption that the errors, and hence the dependent variable y conditional on the explanatory variables, are normally distributed, was tested for. Tests for heteroscedasticity of the residuals and multicollinearity between the explanatory variables were performed. The model’s specifications, e.g. omitted variable bias or the linearity assumption, were also tested. All of the tests support that the assumptions of the model hold.

4.1 Univariate and bivariate analysis of the variables

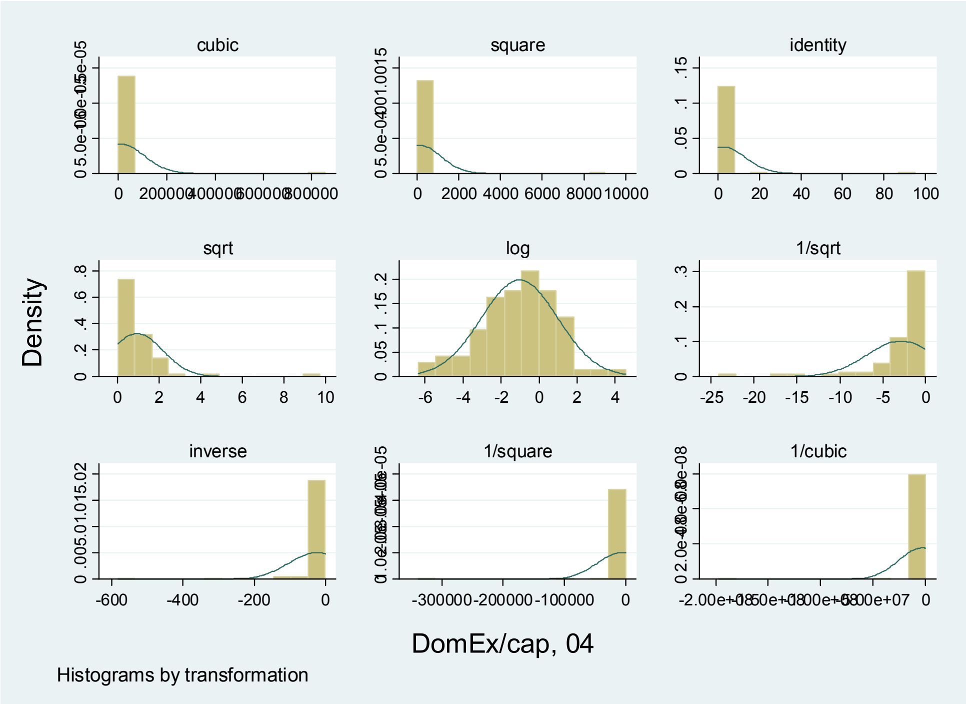
The dependent variable was transformed; the estimation of the regression coefficients do not require normally distributed residuals, but the residuals need to be normally distributed for the t-tests to be valid. A common cause of non-normally distributed residuals is non-normally distributed outcome and/or predictor variables. The summary statistics shows that our dependent variable, namely the public domestic expenditures on HIV and AIDS is not normally distributed marginally. The mean exceeded than the median (table 3), which indicates a positive skewness. The kurtosis was much greater than 3 which is the kurtosis of a normal (Gaussian) distribution, which indicates that the distribution has heavier tails.

Table 3 Descriptive statistics for the domestic public expenditures per capita, 2004

domestic public expenditures per capita, 2004				
	Percentiles	Smallest		
1%	.0017179	.0017179		
5%	.0098273	.0030837		
10%	.0323192	.0043663	Obs	81
25%	.1039845	.0072835		
50%	.4153908		Mean	2.403252
		Largest	Std. Dev.	10.66607
75%	1.565415	4.786399		
90%	3.851129	7.776968	Variance	113.7651
95%	4.524433	17.02529	Skewness	8.25474
99%	94.98671	94.98671	Kurtosis	71.89583

Transforming the dependent variable is useful in order to reduce or remove the heteroscedasticity of the residuals. In order to reduce a positive skewness, transforming the variable by e.g. the square root, the logarithm, the inverse are common methods. Testing this using Stata shows that the method giving the best approximation to the normal distribution is to transform the variable with the natural logarithm:

Figure 4 Distribution of the domestic public expenditures related to HIV and AIDS compared to the normal distribution after transformation by selected methods.



Transforming the dependent variable logarithmically significantly reduces the skewness, but the distribution now has a slight negative skewness (table 4). The kurtosis is now closer to three.

Table 4 Descriptive statistics for the dependent variable, the logarithm of the domestic public expenditures per capita, 2004

Logarithm of the domestic public expenditures per capita, 2004				
	Percentiles	Smallest		
1%	-6.36666	-6.36666		
5%	-4.622586	-5.781615		
10%	-3.432093	-5.43385	Obs	81
25%	-2.263513	-4.92215		
50%	-.8785356		Mean	-1.030097
		Largest	Std. Dev.	2.01471
75%	.4481509	1.565778		
90%	1.348366	2.051167	Variance	4.059056
95%	1.509492	2.8347	Skewness	-.2677454
99%	4.553737	4.553737	Kurtosis	3.220231

A simple linear regression was performed for each independent variable and the domestic public expenditure per capita to determine which explanatory variables are to be included in the final multiple regression analysis. For an explanatory variable to be considered a candidate for the multiple regression the p-value should be less than 20%. However, as stated above, the variable not satisfying this requirement might still be included in further analysis if it is considered important conceptually.

Descriptive statistics and associations

In the further analysis, GDP per capita was used instead of GNI per capita since the correlation was slightly better (table 5).

Table 5 Correlation matrix shows higher correlation between HIV expenditures and GDP per capita than for GNI per capita, both with and without logarithmic transformation

	lndomex	domexcap	GDP/cap	GNI/cap
GDP/cap	0.5129	0.5875	1	
GNI/cap	0.5086	0.5847	0.9982	1

The correlation for all variables[5] with and without transformation of the dependent variable is found in table 6. The number of observations included in the correlations is in the row below.

Table 6 Correlation coefficient between the domestic public expenditures on HIV and AIDS and its potential explanatory variables

	lndomex	domexcap
GDP per capita	0.5237	0.2524
	80	80
Budget support per capita	0.1093	-0.0074
	81	81
Total debt service (% of GNI)	0.1781	-0.141
	78	78
Tax revenue (% of GDP)	0.1624	0.2676
	46	46
Number of hospital beds (per 10 000 population)	0.1919	0.1809
	50	50
Nurses (density per 1 000 population)	0.1588	0.0557
	80	80
Physicians (density per 1 000 population)	0.0875	-0.0499
	81	81
Literacy rate	-0.0111	0.2309
	81	81
Political stability	0.5634	0.2579
	80	80

[5] The cross correlation coefficients are in the annex, table A3

Corruption Index	-0.0062	-0.0046
	61	61
Government expenditure on health at average exchange rate (US$) as % of total health expenditures	0.3601	0.0927
	81	81
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of domestic public health expenditures	-0.2609	-0.0528
	81	81
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of people living with HIV	-0.2515	-0.0447
	75	75
People with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%)	0.5336	0.3356
	74	74
Estimated number of people living with HIV, children (0-14)	0.0016	-0.0139
	73	73
Estimated HIV prevalence, children (0-14)	0.2566	0.3109
	73	73
Estimated number of people living with HIV, adult and children	0.0262	0.0095
	76	76
Estimated number of people living with HIV, adults (15+)	0.032	0.0101
	75	75
HIV prevalence, adult (14-49)	0.3444	0.4468
	76	76
number of people in need of ART	0.0688	0.0134
	73	73

The correlation between ODA as percentage of government health expenditures of domestic public health expenditures and expenditures related to HIV and AIDS is in fact negative. That could support the suspicion that governments receiving a lot of aid for HIV and AIDS prioritize other budget posts. The scatter plot in figure 5 confirms that ODA as percentage of government health expenditures shows a tendency of a negative trend, but only after a certain threshold. When the ODA is under 60 % of the health expenditures, no pattern is observable (figure 6). When the ODA exceeds 60% of the health expenditures, there is an almost perfect linear decline in public expenditures for HIV and AIDS. The observations are too scarce to affirm this tendency (n=5), this should be further investigated when more data is available on this matter.

Figure 5 The Official Development Assistance plotted against the domestic HIV expenditures

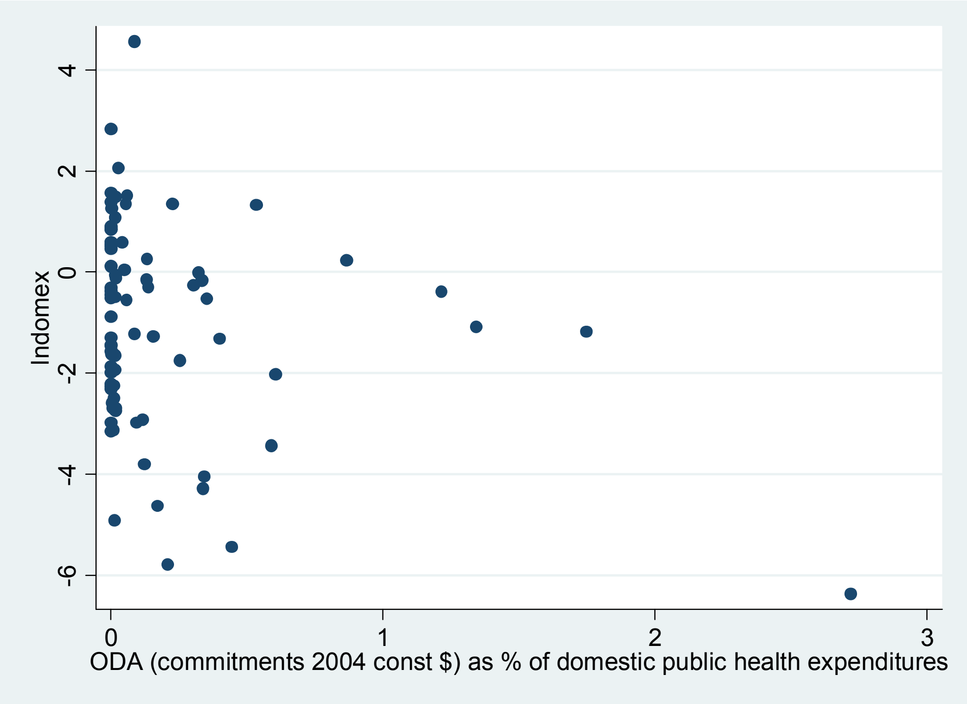
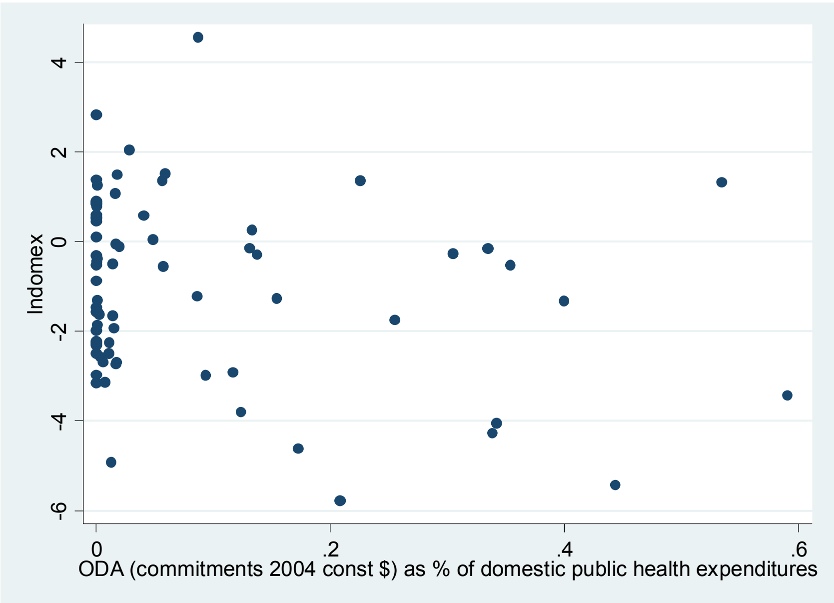


Figure 6 The share of Official Development Assistance plotted against the domestic HIV expenditures when the share is 60% or less



Debt service as percentage of GNI has a negative correlation with domestic public expenditures related to HIV and AIDS when the latter is untransformed. This supports our a priori assumption that countries which use a greater share of the money available to pay back debt would use less on HIV and AIDS. However, with the logarithmic transformation of the dependent variable the correlation becomes positive. The scatter plot in figure 7 shows that Botswana is a prominent outlier; when this observation and the observation of Barbados are deleted there is no obvious pattern.

Figure 7 Debt service as % of GNI plotted against HIV expenditures without transformation

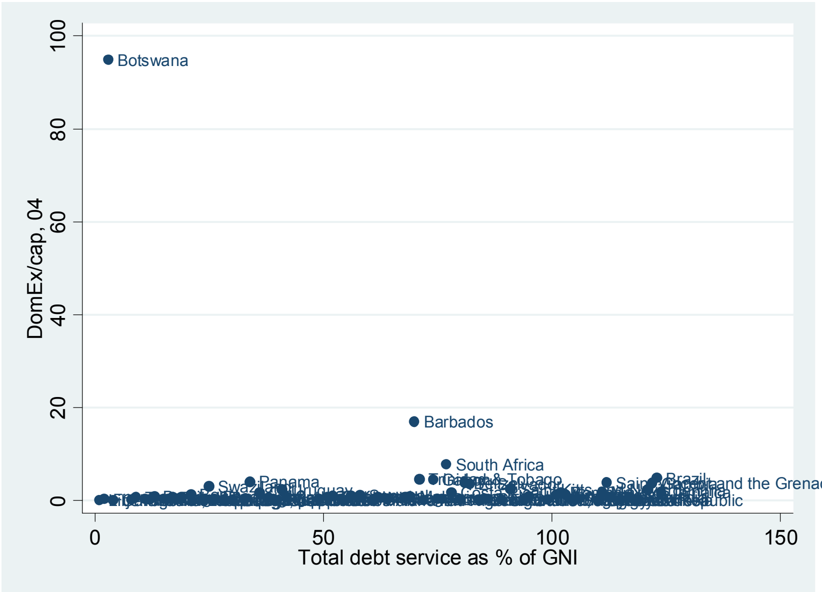
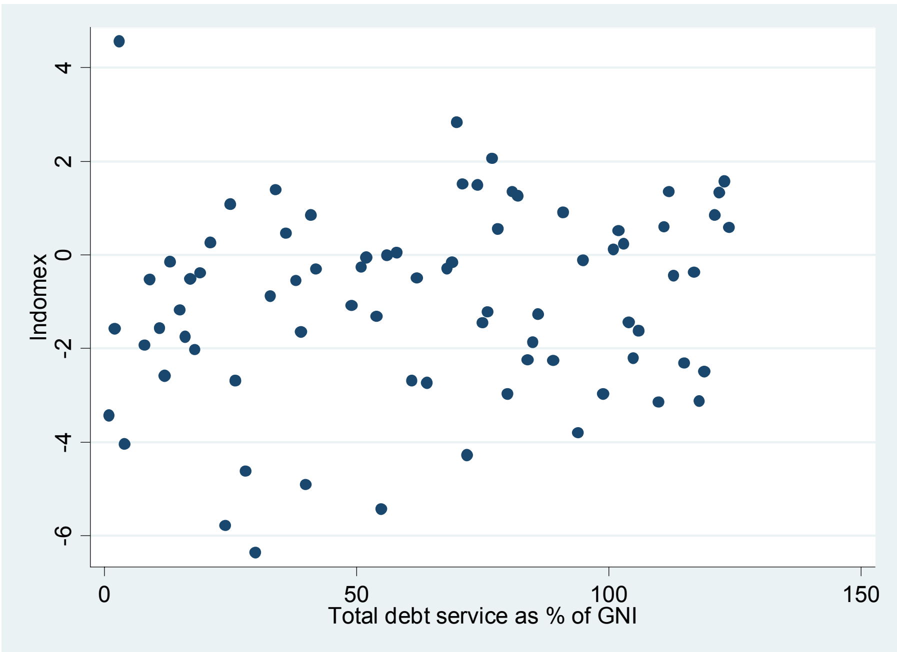


Figure 8 Debt service as % of GNI plotted against the logarithm of HIV expenditures



After the variable for GDP per capita was logarithmically transformed, the scatter plot shows a clearly nonlinear relationship between domestic public expenditures on HIV and AIDS per capita and the GDP per capita (figure A1). The same is indicated by a test of linearity with the final model with no transformation of GDP (figure A2). The plots provide evidence of a concave relationship between GDP and HIV spending. A way to come around this is to introduce a variable of GDP squared, or simply to do a logarithmic conversion of the GDP, hence avoiding the decrease the degrees of freedom and easing the interpretation of the coefficient. The explanation is that HIV expenditures rise with rising GDP, but at a decreasing rate (figure A1). The logarithmic transformation of the variable GDP also changes the interpretation of the coefficient in the regression model. When we have log of GDP, it is interpretable as an X% change in GDP being associated with a Y% change in HIV expenditures, while for GDP it is a $X change in GDP being associated with a Y% change in HIV expenditures.

Simple regression

The criterion for the inclusion of the variables to be considered in the multiple regression was their significant t- statistic and thus a p-value of around 20% or less when doing a simple regression. If, however, the variable is considered important enough from the conceptual framework, it might still be included in further analysis even if it does not satisfy these statistical criteria. Twenty-one variables were tested. Table 7 provides the results from the simple regressions.

Table 7 results from the simple regressions

Proxy variables	logarithm of domestic public expenditures on HIV and AIDS				
	coeff	standard error	t-value	p-value	r-square
GDP	0.0004	0.0001	5.43	0.000	0.274
Logarithm of GDP	0.9712	0.1611	6.03	0.000	0.318
Budget support per capita	0.0035	0.0036	0.98	0.332	0.012
Total debt service (% of GNI)	0.0096	0.0061	1.58	0.119	0.032
Tax revenue (% of GDP)	0.0350	0.0320	1.09	0.281	0.026
Number of hospital beds (per 10 000 inhabitants)	0.0141	0.0104	1.35	0.182	0.037
Nurses (density per 1 000 inhabitants)	0.1392	0.0980	1.42	0.160	0.025
Physicians (density per 1 000 inhabitants)	0.1536	0.1968	0.78	0.437	0.008
Literacy rate	-0.0009	0.0093	-0.1	0.921	0.000

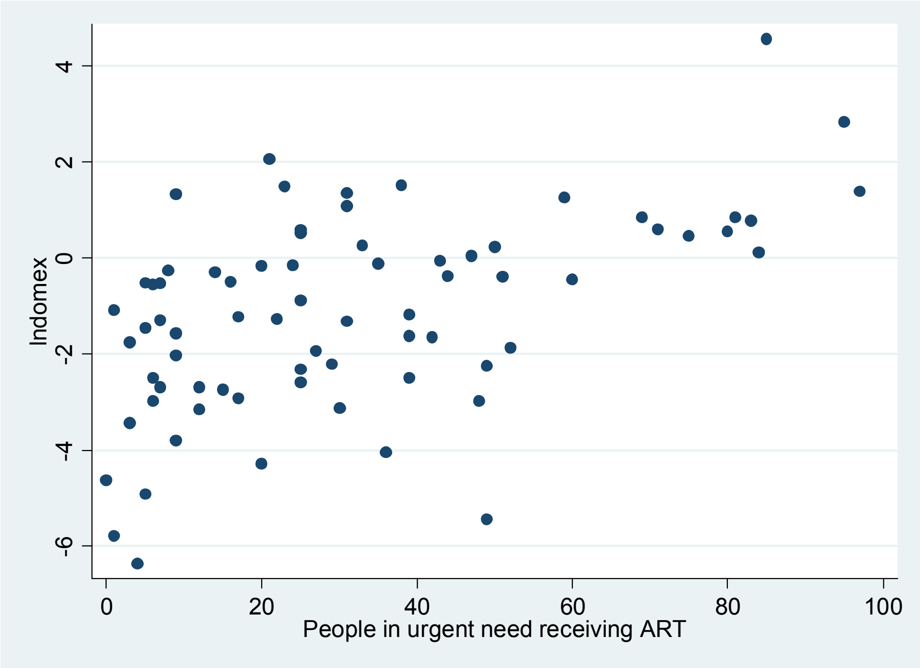
Political stability	1.3673	0.2270	6.02	0.000	0.318
Corruption Index	-0.0093	0.1943	-0.05	0.962	0.000
Government expenditure on health at average exchange rate (US$) as % of total health expenditures	3.8775	1.1301	3.43	0.001	0.130
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of domestic public health expenditures	-1.2514	0.5210	-2.4	0.019	0.068
ODA Commitments for HIV and AIDS in 2004 (2004 in constant dollars) as a ratio of people living with HIV	-0.0010	0.0004	-2.22	0.029	0.063
People with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%)	0.0431	0.0080	5.35	0.000	0.285
Estimated number of people living with HIV, children (0-14)	0.0000	0.0000	0.01	0.990	0.000
Estimated HIV prevalence, children (0-14)	191.2753	85.4895	2.24	0.028	0.066
Estimated number of people living with HIV, adult and children	0.0000	0.0000	0.23	0.820	0.001
Estimated number of people living with HIV, adults (15+)	0.0000	0.0000	0.27	0.785	0.001
HIV prevalence, adult (14-49)	0.1077	0.0341	3.16	0.002	0.119
number of people in need of ART	0.0000	0.0000	0.58	0.563	0.005

The regression analysis gives a p-value below the predetermined critical value of 20% for the following variables: GDP per capita, total debt service (% of GNI), number of hospital beds (per 10 000 population), nurses (density per 1 000 population), political stability, government expenditure on health at average exchange rate (USD) as % of total health expenditures, ODA Commitments for HIV and AIDS as a ratio of domestic public health expenditures, ODA Commitments for HIV and AIDS as a ratio of people living with HIV and AIDS, people with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%), and adult HIV prevalence (14-49 years). Budget support per capita, Tax Revenue (% of GDP), and the estimated HIV prevalence (0-14 years) was decided to be included in search for the final model, since they are important conceptually. However, Tax Revenue was only available for 46 of the countries where also information on HIV and AIDS expenditures is available. Since Stata automatically deletes the observations with any explanatory variables missing, inclusion of the Tax Revenue would have deleted too many countries from the data set available for regression analysis. The literacy rate was also included in spite of its large p-value since it showed significance when the dependent variable was untransformed.

In the search for the final model, ODA Commitments for HIV and AIDS is represented as a ratio of domestic public health expenditures instead of as a ratio of people living with HIV and AIDS since the correlation with HIV spending was slightly better (table 6). The five greatest values of the variable of ODA Commitments for HIV and AIDS in 2004 as a ratio of domestic public health expenditures were deleted since, as explained earlier, they were the ones creating the negative association with the dependent variable.

The variable of people with advanced HIV infection receiving antiretroviral combination therapy (%) (ART coverage), showed a strong correlation with the HIV and AIDS expenditures (figure 9), and has a p-value of under 0.001, but this variable causes an endogeneity problem. There is a high risk of endogeneity since government expenditures for HIV and AIDS is probably increasing the coverage rate of ART. The problem is still present even though the model is only for the year 2004, since the variable of ART coverage is from 2005. If this variable was to be included it should be using an instrumental variable, i.e. a variable highly correlated to ART, but with no endogeneity problems with the spending on AIDS. The coverage of treatment of diabetes or a disease requiring high medical expenditures like cancer, or preferably another disease with stigma attached to it could be an instrument variable. This variable is not included in the search for the final model, but should be explored closer by instrumental variable techniques.

Figure 9 Antiretroviral treatment (ART) coverage plotted against HIV expenditures



Thus, in summary, ten explanatory variables were considered for inclusion in the final model: The logarithm of GDP per capita, total debt service (% of GNI), number of hospital beds (per 10 000 population), nurses (density per 1 000 population), political stability, government expenditure on health as % of total health expenditures, ODA Commitments for HIV and AIDS in 2004 as a ratio of domestic public health expenditures, adult HIV prevalence (14-49 years), budget support per capita, and the estimated HIV prevalence (0-14 years) per capita.

4.2 The regression model

The multiple regression model is estimated by the Ordinary Least Squares (OLS) method. It analyzes what determines the level of domestic public expenditure, how a change in the explanatory variables, i.e. the GDP level, prevalence etc. would affect the level of public expenditure. This is the model used to find the coefficients of equation (iii) from the conceptual framework in chapter 3.2, and thereby allowing us to estimate the domestic public expenditures on HIV and AIDS, predicting in particular for countries for which this variable is unobserved. The simplified representation of (iii) is thus:

$$y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + e$$

The dependent variable y_i denotes the logarithm of the domestic public expenditures on HIV and AIDS per capita and β_0 is the constant term. The explanatory variables are denoted by $x_1,...,x_k$. A dummy variable to examine for regional patterns, taking the form of shifts in the constant term, will be introduced. The *e* is the error term; it represents all variations in the dependent variable not explained by the model. The coefficients β_i are unknown parameters. The parameter measures the effect of a change in its explanatory variable *ceteris paribus* (all other variables constant) on the dependent variable. That is if for example β_2 equals 5, a 1 unit increase in the variable x_2 will increase the domestic public expenditure related to HIV and AIDS by 5 percent.

Some assumptions of the random errors need to be fulfilled for the estimators to be Best Linear Unbiased Estimator (BLUE) according to Gauss-Markov theorem:

1. The expectation of the error term is zero
2. The variance of the error term is constant and they are uncorrelated
3. Any pair of errors is uncorrelated
4. The random errors are normally distributed
5. The values of the explanatory variables are not random; there are no exact linear relationship between them and there are more observations than explanatory variables

These assumptions are consistent with the ones made in chapter 3.2 when deriving equations (i)–(iii).

When we have non-experimental data, these assumptions might often not be fulfilled, and we might not have BLUE estimates, but they can still be consistent or unbiased.

4.3 Searching for the optimal model

The model was estimated doing forward regression, since Stata automatically deletes the observations with any missing variables. The forward regression was done adding one by one the variables with the highest correlations (see table A3), and the choice was made to eliminate those variables for which a t-test gave a high p-value when included in the model. Table A4 in the annex gives the forward regression with the significant variables. The forward regression yielded a model with GDP per capita, budget support per capita, political stability, and HIV prevalence as explanatory variables (regressors). The correlation matrix of the variables included in the final model is given in table 8 with and without log-transformation of the dependent variable.

Table 8 Correlation between the variables included in the multiple regression model and their correlation with the unconverted domestic expenditures on HIV and AIDS, n=74

	lndomex	domexcap	Political stability	Logarithm of GDP/cap	GDP per capita	HIV prevalence, adult	Budget support per capita
lndomex	1						
domexcap	0.4444	1					
Political stability	0.5586	0.2868	1				
Logarithm of GDP/cap	0.5431	0.2377	0.4857	1			
GDP per capita	0.5043	0.2699	0.5288	0.8665	1		
HIV prevalence, adult (14-49)	0.3446	0.4472	0.12	-0.0979	-0.0552	1	
Budget support per capita	0.0441	-0.0814	0.0046	-0.2623	-0.2676	0.0249	1

Interestingly, political stability is the variable which is most strongly correlated with the public spending on HIV and AIDS, followed by GDP which was expected to have a high correlation (table 9). These two variables also have a high correlation, 0.49, posing a potential multicollinearity problem. However, how serious this will be depends on the sample size, and the diagnostic gave evidence that this was not a problem. Since political stability is survey based, there is no chance of GDP to be a component of the political stability indicator. The correlation of the two is probably present because poor countries are often among the most unstable. The correlation of the budget support is not very strong, and in fact negative when public spending is untransformed.

The budget support per capita was retained, with a t-test yielding a p-value of 0.052[6] (table 9).

[6] We expected this variable to have a positive association with HIV and AIDS spending, thus it is not necessary to check both tails of the distribution. By default the t test probabilities are two sided, for one-side tests, the p-values could actually be divided in half.

A mentioned in section 3.3, the political stability was chosen over six dimensions of governance by the World Bank. The reason was the six dimensions; voice and accountability; political stability; government effectiveness; regulatory quality; rule of law; and control of corruption, are highly collinear, and that the political stability had the highest correlation with the dependent variable. However, the five other dimensions of governance were added to the current model to perform a joint hypothesis test. The joint hypothesis tests were done with the subset of coefficients, which is usually useful when we have several conceptually related predictors as in this case. The result when testing by an F-test the joint hypothesis that all of the six governance indicator's coefficients are equal zero gave rejection at a p-value of 0.0223. When we test the hypothesis that all governance indicators except the political stability are all equal zero at the same time, the p-value is 0.2246 (table A5), which means that we cannot reject this hypothesis. These test and the fact that the dimensions are highly correlated with each other, indicate that it is sufficient to include only one. Political stability is the one that are strongest correlated with domestic public expenditures for HIV and AIDS, and the only one that is statistically significant in the current model. Altogether, the choice of political stability as a measurement for governance is supported, and the model is not changed.

The next step to identifying the optimal final model was to test for regional variations and interactions between the explanatory variables. The residuals of the model without dummies were examined to check for patterns where the model most wrongly predicts the expenditures[7], and then dummies are introduced to see if the model can be improved. In the residual plot for the model without dummies, the residuals seemed symmetrically distributed around zero which is consistent with the normal errors assumption. However, listing the ten highest and lowest residuals and their regions indicated that a regional pattern is indeed present. Among the ten largest residuals, seven were for Sub-Saharan countries, and three for Latin America. According to the analysis of the residuals, the model explain the least the low income countries and the countries in Sub Saharan Africa and Latin America, giving reason to create a dummy variable for the two regions. Including the dummies for Latin America (reg4) and Sub Saharan Africa (reg8) increases the adjusted r-square and thus the explanatory power of the model, and the dummies are significant (table A8). The two dummies together gave the highest r-square, and it shows that Sub Saharan Africa and Latin

[7] Residual plots and tables are in the annex, figure A3, and table A6-A7

America spends more per capita than others with the same levels of GDP, budget support, political stability and prevalence.

The interaction between variables was tested; two explanatory variables are said to interact if the effect of one explanatory variable on the response variable depends on the level of the second explanatory variable, i.e. the first derivatives are not constant. Sub Saharan Africa was suspected to behave differently, and the scatter plot of expenditure on HIV and AIDS versus GPD level showed a flatter slope than for other regions (figure A4). Thus, an interaction was tested between the level of GDP and the region (reg8gdp). The flatter slope indicates that for the same income increase, less of the money goes to expenditure on HIV and AIDS, However, the intercept dummy shows a higher level of spending on average for the region. The result is confirmed by the significance of the interaction variable. For low levels of GDP, Sub Saharan Africa spends more, but the slope is not as steep as for other regions so when GDP reaches higher levels, the spending is not that high. Table 9 shows the results from the models with and without dummies:

Table 9 Regression statistics with and without dummy variables, coefficients with standard errors in the parenthesis.

	lndomex	lndomex
Logarithm of GDP/cap	0.848***	1.256***
	(0.173)	(0.244)
Budget support per capita	0.052#	0.059*
	(0.027)	(0.024)
HIV prevalence adult 15-49 years	0.111***	0.095**
	(0.026)	(0.029)
Political stability	0.736**	0.587*
	(0.251)	(0.227)
Sub-Saharan Africa	--	5.197*
		(2.148)
Latin America	--	1.301**
		(0.407)
reg8gdp	--	-0.575#
		(0.319)
Intercept	-9.564***	-12.948***
	(1.094)	(1.656)
N	74	74
R2	0.549	0.664

Adjusted R2	0.522	0.628
*** p<0.001; ** p<0.01; * p<0.05; # p<0.1; two tailed		

The latter model had a better adjusted r- square, and thus a better explanatory power, and seemed more plausible. Hence it was chosen as the preferred model.

4.4 The final model

The final model explains the domestic public expenditures for HIV and AIDS by means of four variables; GDP per capita, budget support per capita, political stability, and HIV prevalence.

$$\ln domex = -12.948 + 1.301\, reg4 + 5.197\, reg8 + 1.256 \ln gdp - 0.575\, reg8 * \ln gdp + 0.059\, budgetcap + 0.095\, prevalence + 0.587\, polstab + e$$

Since the model includes several dummies, it is easier to see its implication by displaying the equation for the three regions separated.

For Sub-Saharan Africa the equation will be:

$$\ln domex = -7.751 + 0.681 \ln gdp + 0.059\, budgetcap + 0.095\, prev + 0.587\, polstab + e$$

For Latin America the equation will be:

$$\ln domex = -11.647 . + 1.256 \ln gdp + 0.059\, budgetcap + 0.095\, prev + 0.587\, polstab + e$$

For the rest, the equation is:

$$\ln domex = -12.948 . + 1.256 \ln gdp + 0.059\, budgetcap + 0.095\, prev + 0.587\, polstab + e$$

When the equations are displayed separately, the atypical pattern for Sub Saharan Africa is more apparent. GDP per capita is not driving expenditures as much as in other regions, in fact the effect is halved. An increase of 1% in the GDP per capita gives 1.26% increase in expenditures in other regions, but only 0.6% increase for Sub Saharan Africa. Initially this region has, however, higher expenditures per capita reflected by the higher constant.

As the model is now, it includes the three conceptual elements; the amount of money to spend, the ability to spend the money, and the need for money. No variables representing the willingness to spend money were found, and where available, they should be included in the model.

The GDP per capita together with the budget support shows that the more money available, the more the government spends on HIV and AIDS (table 10 gives their mean values by region). When removing the budget support, the other coefficients stay more or less the same, so this proxy accounts for the fact that there are countries with a low GDP still being able to spend money because of the received funding.

Table 10 Mean values by region for GDP and budget support per capita

Region		GDP/cap (curUS$)	Budget support per capita
Caribbean	CAR	5,043	7.51
East Asia	EA	1,066	0.00
Eastern Europe and Central Asia	EECA	2,869	1.19
Latin America	LA	3,123	2.04
Middle East & North Africa	MENA	3,353	0.64
Oceania	OCE	3,148	113.62
South & South East Asia	SEA	1,307	0.73
Sub-Saharan Africa	SSA	1,052	3.78
Western and Central Europe	WCE	6,111	4.79
Total		2,534	6.85

The ability to spend money is captured by the variable political stability. From the results in table 9 we see that a one unit increase in the political stability gives a 0.59% increase in the spending. The political stability reflects the statistical compilation of responses on the quality of governance, asking questions like “How does risk of political violence influence government?” It is confirmed that this variable is indeed of importance when it comes to expenditures on HIV and AIDS. Sub Saharan Africa has a lower political stability than the total mean, while Caribbean countries are much more politically stable than the average (table 11).

Table 11 Mean values by region for political stability

Region	Political Stability
CAR	0.471
EA	0.205
EECA	-0.350
LA	-0.313
MENA	-0.864
OCE	0.363
SEA	-0.796
SSA	-0.581
WCE	-0.036
Total	-0.386

Prevalence is indicating the burden of the disease in the country. Table 12 shows clearly that Sub-Saharan Africa has a higher prevalence than any of the other regions.

Table 12 Mean values by region for HIV prevalence among adults (15 to 49 years)

Region	HIV prevalence
CAR	1.661
EA	0.033
EECA	0.364
LA	0.649
MENA	0.201
OCE	0.900
SEA	0.387
SSA	6.915
WCE	0.070
Total	2.679

Per capita values are used to adjust for the size of the population; Mexico spends about 185 million dollars on HIV and AIDS which is a lot more than El Salvador's 23 millions, but adjusted for population size El Salvador outranks Mexico by spending almost one and a half dollar more per capita (3,5 USD/cap compared to 1.8).

Table 13 Mean values for domestic public expenditures on HIV and AIDS for the 81 observations

Region	DomEx/cap,	DomEx04
CAR	4.418	2 402 797
EA	0.076	97 991 976
EECA	0.382	10 107 024
LA	1.799	77 712 159
MENA	0.124	3 946 111
OCE	0.920	78 821
SEA	0.133	15 731 257

SSA	4.149	23 453 889
WCE	0.413	3 466 050
Total	2.403	29 665 080

4.5 Diagnostics for the model

Identifying unusual and influential variables

The model was checked for outliers and influential observations. In linear regression, an outlier is defined as an observation with a large residual[8], and the studentized residuals were checked for[9]. A large squared residual indicates that an observation has large differences in predicted and actual expenditures on HIV and AIDS.

Leverage points[10], when an observation has an extreme value on a predictor variable, can have an effect on the estimate of the regression coefficients. Leverage over (2k+2)/n, where k is the number of independent variables and n the number of observations e.g. 0.2162 in this case, should be examined closer.

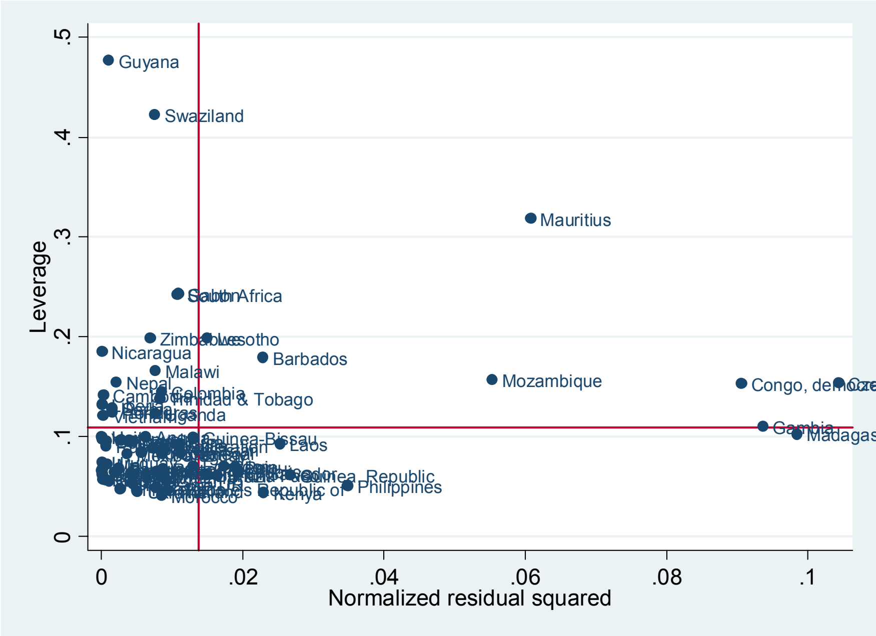
As a synthesis of the tests performed, figure 10 gives the leverage-versus-squared-residuals plot. The lines mark the means of leverage (horizontal line) and squared residuals (vertical line).

[8] See tables in the annex; A9-A10

[9] A Studentized residual, named in honor of William Sealey Gosset, who wrote under the pseudonym Student, is a residual adjusted by dividing it by an estimate of its standard deviation. Studentization of residuals is an important technique in the detection of outliers.

[10] See table A11-A12 in the annex

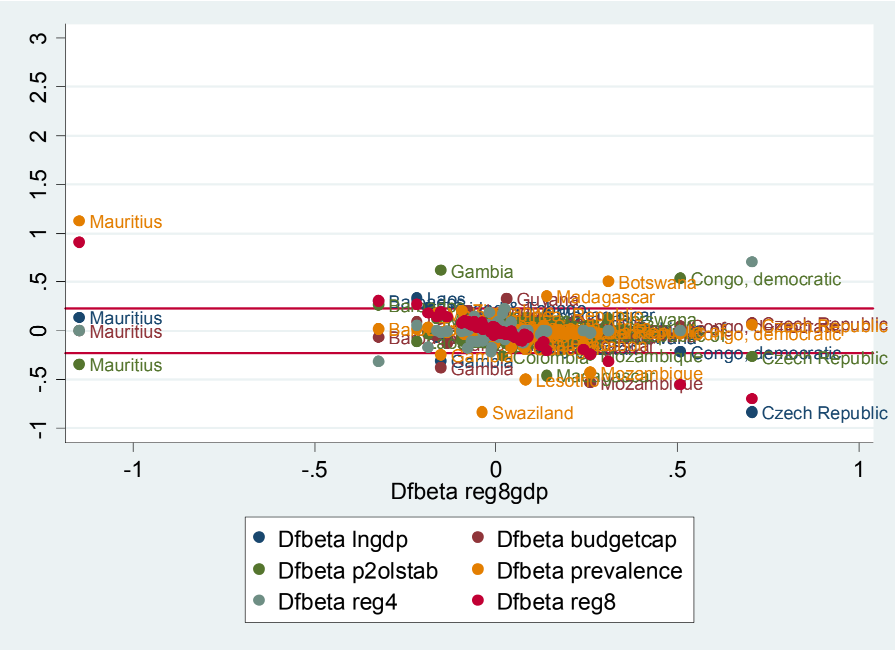
Figure 10 Leverage versus squared residual plot to identify the most influential observations



Guyana and Swaziland have the greatest leverage, but the model fits the observations very well. This is not necessarily a virtue, since high leverage points can control the regression coefficients to fit them. The Democratic Republic of Congo, the Czech Republic, Gambia and Madagascar have poor fits, but have less influence. Mauritius stands out as an observation that is both ill fit and potentially influential. The actual observation on Mauritius is 0.1 USD per capita in 2004, which is below average. The estimation for Mauritius is a lot higher, meaning that the model predicted Mauritius to spend more than it did. Since their HIV prevalence level is around the average with a value of 0.42%, this is probably driven by Mauritius being an upper middle-income country with a GDP per capita in the 75th percentile, and with a high degree of political stability.

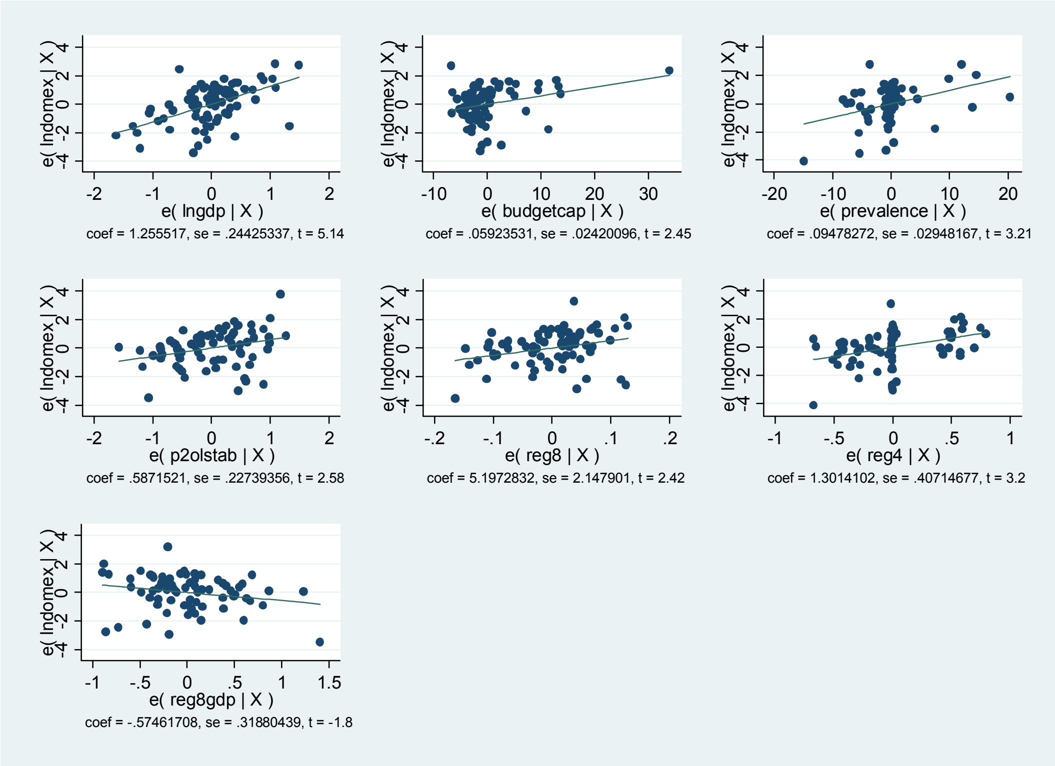
Influential observations can also be measured by specific statistics measuring how much each coefficient will change with the removal of a single observation. One such measure is DFBETA, which indicates how many standard errors a coefficient would change if observation i is deleted. Values in excess of $2/\sqrt{n}$, which is 0.232 with 74 observations, merits further investigation. In figure 11, the two horizontal lines indicate this as the cut off line, and the influence of Mauritius is again obvious. Now it is evident that Mauritius is the most influential through the variable reg8, the dummy of Sub Saharan Africa, since Mauritius is an upper middle income country. This could mean that the model is not as good in predicting expenditures of higher income countries in Sub Saharan Africa. Swaziland is influencing the coefficient of HIV prevalence; with a prevalence of 33% it is the country with the highest prevalence and far from the mean of 0.68%. Czech Republic is influencing the GDP per capita coefficient, and the Democratic Republic of Congo and Gambia are influencing the coefficient of political stability, the latter being more political stable than average, and the former less.

Figure 11 The two horizontal lines show values in excess of $2/\sqrt{74}$ indicating influential observations



To check how one observation influences the regression line, the added variable plots (also known as partial regression leverage plot or adjusted variable plot) were examined (figure 12). They depict the relationship between the dependent variable with one of the independent variables, adjusted for the effect of the other independent variables. This helps to uncover observations exerting a disproportionate influence of the estimated regression model. Guyana, with a budget support per capita of 38.6, is showing to be influencing this variable. If this observation is deleted, there is almost no change in the coefficient.

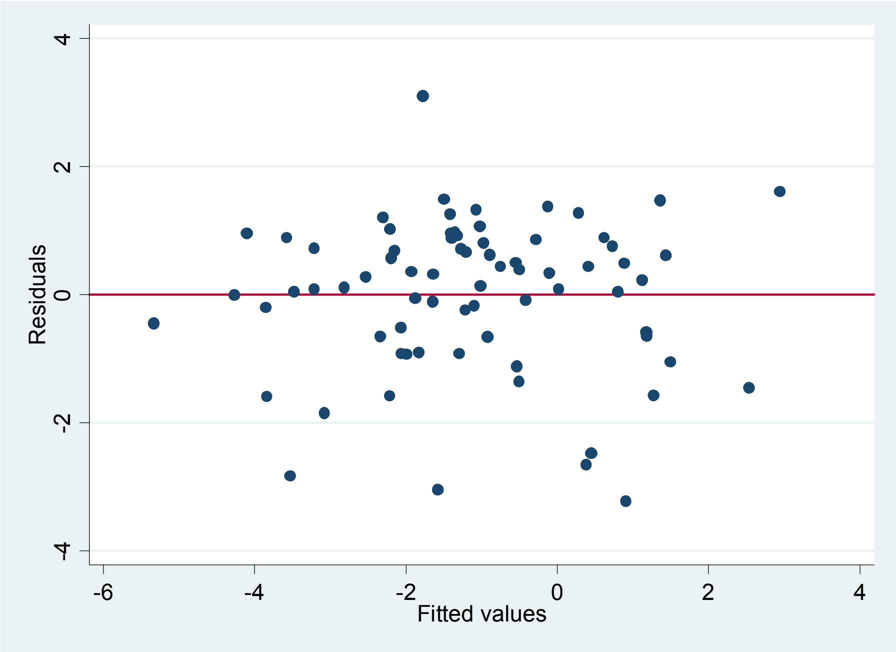
Figure 12 The added variable plots for all explanatory variables in the final model



Checking normality of errors

The normality of the residuals is required for valid hypothesis testing, but usually it is not required to obtain unbiased coefficients of the regression. The residual plot (figure 13) shows residuals that seem symmetrically distributed around zero. A symmetric density is necessary for normality to hold, but not sufficient, so the plot is not inconsistent with the normality assumption.

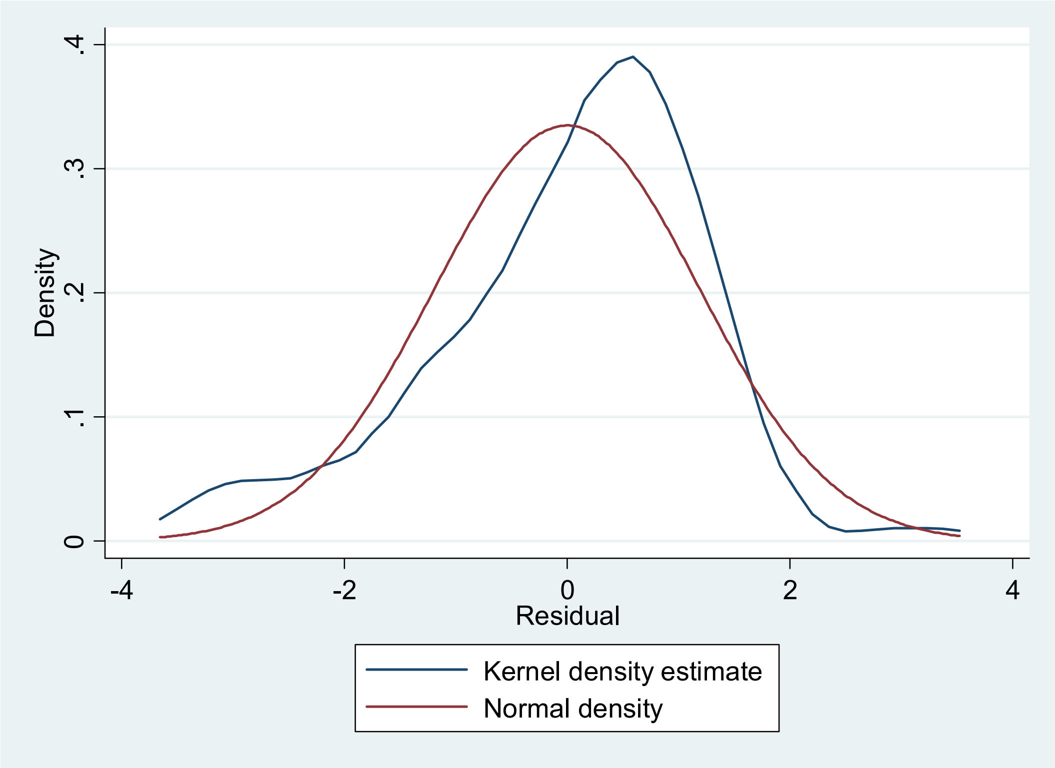
Figure 13 The residual plot



A kernel density estimate[11] was plotted for the model overlaid by a normal density plot (figure 14). This plot seems less coherent with the normality of the errors, the kurtosis is almost 4, and the distribution is positively skewed. However, the deviations do not seem too substantial.

[11] A standardized normal probability plot, and a plot with the quantiles of the residuals plotted with the quantiles of the normal distribution is found in the annex (figure A5 and A6)

Figure 14 The kernel density estimate overlaid by a normal density plot



The Shapiro-Wilks test of the null hypothesis that the residuals are normally distributed gave a p-value of about 0.002, which shows that we will have to reject the null hypothesis of normally distributed errors.

Table 14 Shapiro-Wilks test for normal data

Variable	Obs	Skewness	Kurtosis	z	Prob>z
e1	73	0.93967	3.842	2.934	0.00167

However, a test for symmetry of the distribution based on inter-quartile range[12] gives presence of mild and severe outliers (table 15). The presence of severe outliers should be sufficient evidence to reject normality at a 5% significance level, while moderate outliers are common for any sample size. Table 16 suggests that there are no severe outliers, and the distribution seems fairly symmetric, giving evidence that the errors have an approximate normal distribution.

[12] Written by Lawrence C. Hamilton Department of Sociology, University of New Hampshire, the stata command being “iqr”.

Table 15 Test for inter-quartile range that assuming symmetry of the residual distribution

```
   mean= -5.1e-11          std.dev.=  1.178          (n= 73)
 median=  .2506    pseudo std.dev.=  1.135        (IQR=  1.531)
10 trim=  .1203
                                                low        high
                                                -------------------
                                inner fences    -2.952       3.171
                           # mild outliers      3           0
                           % mild outliers      4.11%       0.00%

                                outer fences    -5.249       5.468
                           # severe outliers    0           0
                           % severe outliers    0.00%       0.00%
```

Checking for heteroscedasticity of residuals

Heteroscedasticity of the disturbances was tested with a White test (table A13) and a Breusch-Pagan / Cook-Weisberg test (table A14). They test the assumption of a constant error variance by examining whether squared standardized residuals are linearly related to predicted *y*. Both test the null hypothesis that the variance of the residuals is homogenous, thus a small p-value means that the null hypothesis must be rejected and accept the alternative hypothesis that the variance is heterogeneous. The White test has a p-value of 0.0421, and Breusch-Pagan a more robust p-value of 0.3768. We do not reject the null hypothesis in the latter case, suggesting that heteroscedasticity is not a big problem. The tests are very sensitive to the model's assumptions such as the assumption for normality. Therefore they are often used together with the residual plot. Both tests and the residual plot indicate heteroscedasticity is not a problem in this model.

Checking for multicollinearity

When there is a perfect linear relationship among the regressors, the estimates for a regression model cannot be uniquely computed. When two variables move together in a systematic way, that is when they are near perfect linear combinations of one another, the variables are said to be collinear, and when more than two variables are involved the term which defines the movement is multicollinearity (Hill et al., 2001:189). The primary concern is that as the degree of multicollinearity increases, the regression model's estimates of the coefficients become unstable and the standard errors for the coefficients can get very inflated.

The *variance inflation factor* VIF (Stata, 2005) was used to check for collinearity (table 16). As a rule of thumb, a variable whose VIF values are greater than 10 may merit further investigation. Tolerance is defined as 1/VIF, a tolerance value lower than 0.1 is comparable to a VIF of 10.

The only problem exists between the dummy variable for Sub-Saharan Africa and the interaction term reg8gdp. It could be that they capture the same effect to some extent, but they both have high t-values in the model, so they are both kept. Otherwise, there are no indications of multicollinearity, and the correlation between political stability and the GDP per capita does not affect the model.

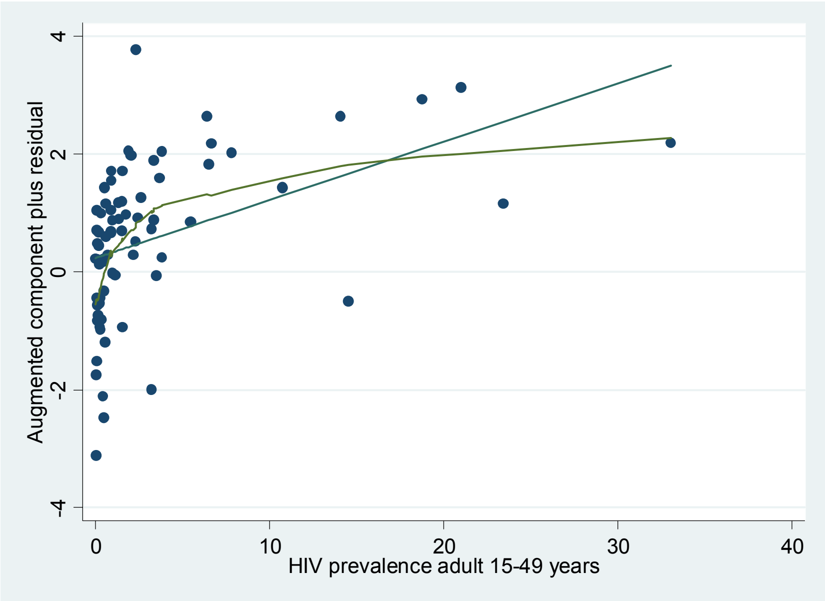
Table 16 Variance inflation factor (VIF) to check for multicollinearity, VIF values should not be greater than 10.

Variable	VIF	1/VIF
-----------------------	------------------	
Sub-Saharan Africa	52.95	0.018885
reg8gdp	46.96	0.021294
Logarithm of GDP/cap	3.74	0.267658
HIV prevalence, adult (14-49)	1.58	0.634773
Political stability	1.43	0.701288
Latin America	1.38	0.725110
Budget support per capita	1.12	0.895731
-----------------------	------------------	
Mean VIF	15.59	

Checking the specifications of the model

In linear regressions the relationship between the dependent and the independent variable is assumed to be linear. In a multiple regression model this can be tested by graphing an augmented component-plus-residual plot, also known as an augmented partial residual plot. There were no great concerns in these plots, after the logarithmic transformation of GDP per capita, but HIV prevalence shows a concave relationship with expenditures rather than a linear one (figure 15). However, the pattern does not seem to fit better than the regression line, and could be influenced by a few observations.

Figure 15 The augmented component-plus-residual plot shows a tendency of unlinearity for the variable of HIV prevalence



The specification of the model was tested by checking the possibility that the model had an omitted variable problem. The Ramsey RESET test (Regression Specification Error Test) is designed to detect omitted variables and incorrect functional form (table 17). The test examines whether the powers of the fitted values are significant when included in the model. Generally the test captures that "if we can significantly improve the model by artificially including powers of the prediction of the model, then the original model must have been inadequate" (Hill et. al, 2001:188). Thus, the rejection of the null hypothesis implies that the original model is inadequate and can be improved. With a p-value of 0.2548, the null hypothesis was not rejected, implying that the test has not been able to detect any misspecifications of the model.

Table 17 Ramsey RESET test using powers of the fitted values of logarithm of the domestic public expenditures on HIV and AIDS

Ho: model has no omitted variables	
F(3, 62) =	1.39
Prob > F =	0.2548

A similar method is performed by the command linktest (table 18), introducing the square of the predictions to see if the coefficient is significant. The predicted values, _hat, should be significant, but the squared predictors, _hatsq, should not if the model is specified correctly.

In this model, the squared predictors were not significant, confirming that omitted variable bias is not a problem.

Table 18 The command "linktest", testing for omitted variable bias

```
      Source |       SS       df       MS              Number of obs =      73
-------------+------------------------------           F(  2,    70) =   64.62
       Model |  178.468441     2  89.2342206           Prob > F      =  0.0000
    Residual |  96.6662522    70  1.38094646           R-squared     =  0.6487
-------------+------------------------------           Adj R-squared =  0.6386
       Total |  275.134693    72  3.82131519           Root MSE      =  1.1751

------------------------------------------------------------------------------
     lndomex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _hat |   .8140786   .1502499     5.42   0.000     .5144147    1.113743
      _hatsq |  -.0691583   .0450902    -1.53   0.130    -.1590879    .0207713
       _cons |   .0426153   .1758232     0.24   0.809    -.3080529    .3932836
------------------------------------------------------------------------------
```

5. The predictions

5.1 Methods

There are different types of countries in the database, creating different problems. Mainly, the prediction problem is addressed in two different ways:

1. The variable to be predicted is in logarithms. Hence when taking the antilog of the predicted logarithmic value, we will on average under-predict the variable of interest. This is because of the properties of the natural logarithm, namely its concave functional form; Jensen's inequality (Sydsæter et al., 1999: section 7.1).
2. For eleven countries, the database does not include information on domestic public expenditures, and predicted values cannot be obtained from the estimated regression model because one or more exogenous variables are missing. These variables will be replaced by regional means and predictions will be compute with these mean values inserted in the estimated regression equation.

The problem with the antilog can be formulated as follows:

Let *n* be the number of countries, Y_i be the observed expenditures on HIV and AIDS, and $W_i = \ln Y_i$ the natural logarithm of the expenditures.

For countries *i* = 1,…, *m*, observations exists for both Y_i and $W_i = \ln(Y_i)$.

For countries *i* = *m*+1,…,*n*, we have unbiased predictions of $W_i = \ln(Y_i)$, but no observations on Y_i and $W_i = \ln(Y_i)$. To construct an unbiased predictor of $Y_i = e^{W_i}$ for countries *i* = *m*+1,…,*n*, we proceed as follows:

Let $\hat{W}_i = \ln(\hat{Y}_i)$ be the predictor of W_i for *i* = *m*+1,…,*n*. It is an unbiased estimator of

$$\mathrm{E}(W_i \mid X_i) = \alpha + \sum_k \beta_k x_{ki} ,$$

where $X_i = (X_{1i}, X_{ki})$.

Furthermore,

$$\mathrm{var}(W_i \mid X_i) = \mathrm{var}(u_i \mid X_i) = \sigma^2$$

Assume that the disturbance is normally distributed, that is $u_i \sim \mathrm{N}(0,\sigma^2)$. Then we know that

$$(W_i \mid X_i) \sim \mathrm{N}(\alpha + \sum_k \beta_k x_{ki}, \sigma^2)$$

$$v_i = e^{u_i} \ \& \ (Y_i \mid X_i) = (e^{w_i} \mid X_i) \text{ are log-normally distributed.}$$

We therefore have (Sydsæter et al., 1999: section 32.13)

$$\mathrm{E}(v_i) = \mathrm{E}(e^{u_i}) = e^{\frac{1}{2}\sigma^2}$$

and

$$\mathrm{E}(W_i \mid X_i) = \mathrm{E}(e^{W_i} \mid XW_i) = e^{\alpha + \sum_k \beta_k x_{ki} + \frac{1}{2}\sigma^2}$$

The conclusion is: Replace $\hat{\alpha}$ by $\hat{\alpha} + \frac{1}{2}\hat{\sigma}^2$ when predicting Y_i from predictions on $W_i = \ln(Y_i)$, that is add half of the estimated disturbance variance estimated from observations for $i = 1,\ldots,m$ to the estimated intercept when predicting the antilog of the log transformed left hand side variable for $i = m+1,\ldots,n$.

To solve this problem, the decision was made to add half of the estimated disturbance variance to the estimated intercept. The mean square error is the residual or error sum of squares divided by the number of degrees of freedom of the sum; it gives an estimate of the error or residual variance.

Since the error term, u_i, is not exactly normally distributed, and therefore $v_i = e^{u_i}$ not exactly log-normally distributed, these new estimates can be further adjusted using one of the properties of the regression predictors. For the countries 1… *m*, that is for the countries with both observations and prediction for the domestic public expenditure on HIV and AIDS available, the predictions and the observations should both sum to the same figure. This is because of the property that the sum of residuals in linear regressions with Ordinary Least Squares always equals zero. That is, had our equations been linear, the problem with the estimates would have not occurred. Thus the suggestion is to adjust the predictions with the previous estimation method by a ratio, k, of the ratio between the sum of the observed and predicted values in countries 1,…, *m*. This way, the sum of the residuals will equal zero and this property would again hold.
That is:

$$\sum_{i=1}^{m} Y_i^* = k * \sum_{i=1}^{m} \hat{Y}_i \Leftrightarrow k = \frac{\sum_{i=1}^{m} Y_i^*}{\sum_{i=1}^{m} \hat{Y}_i}$$

Where Y_i^* is the observed expenditures, and $\hat{Y}_i$ the predictions. The ratio k will then be multiplied with the estimations for countries m+1, …, n thereby adjusting these estimates with the same ratio found with countries 1,…,m. This will give predictions about 30 percent less (k= 0.7036618) than the previous method.

The three-step methods, simply first taking the antilog, next adjusting with the residual variance, and finally, adjusting this again with the property of the residuals (k), give three different results. The antilogs give the lower bound estimates, and the adjustment with the residual variance gives the upper bound. Adjusting with k gives an estimate in between, and in the further analysis of the domestic expenditure on HIV and AIDS, these estimates will be used for the 46 countries.

Predicted values cannot be obtained from the estimated regression model for eleven countries since one or more exogenous variables are missing. The missing values are supposed to be NMAR, i.e. not missing at random (Scheffer, 2002:153); the countries with no data for say, prevalence or GDP are often poor countries with little or no system for collecting statistics. Some of them are also countries in conflict or post-conflict. Two countries, Afghanistan and Iraq, did not even have information on population, and we therefore decided to exclude them from the sample for estimation. For the other nine countries, GDP and prevalence were the only two missing variables (table 19). These missing values were replaced by regional means, even though the regions vary in size. There are other, more elaborate, methods to deal with missing data (Scheffer, 2002), but since the percentage of countries with missing values are so low, only 0.0714%, the method mentioned above was used.

Table 19 Nine countries for which missing values were replaced by regional means (in green)

Region	*countries*	*Logarithm of GDP*	*Budget support*	*HIV prevalence*	*Political Stability*
CAR	Antigua and Barbuda	9.287	0	1.661	1.127
CAR	Cuba	8.188	0	0.07	0.196
CAR	Dominica	8.241	44.045	1.661	1.014
CAR	Saint Lucia	8.448	2.333	1.661	1.448

EA	Democratic Republic of Korea	6.885	0	0	0.044
OCE	American Samoa	7.771	0	0.9	0.722
SSA	Seychelles	9.037	1.485	6.915	0.641
SSA	Somalia	6.236	0	0.87	-2.650
WCE	Albania	7.782	9.602	0.07	-0.7857465

The regression was then performed with these replacements, yielding predictions for these nine countries as well. The same three-step transformations were performed on these predictions as for the countries with complete data sets.

5.2 The predicted expenditures

The 81 observations we had on domestic public expenditures on HIV and AIDS were combined with the 46 predictions from the multiple regression and the nine predictions with missing variables, By summing these values, we obtain an estimated total for 136 low- and middle income countries. The total estimated expenditures for the countries are 2,744 million USD.

The Latin American countries are by far the largest contributors, accounting for about 48% of the domestic public expenditures related to HIV and AIDS (Figure 16). Brazil alone spent about USD 880 millions in 2004. Sub Saharan Africa spends about 28%, i.e. USD 764 millions (figure 17), mainly driven by the upper middle income countries such as South Africa and Botswana.

Figure 16 Total estimated expenditures for 136 low- and middle-income countries

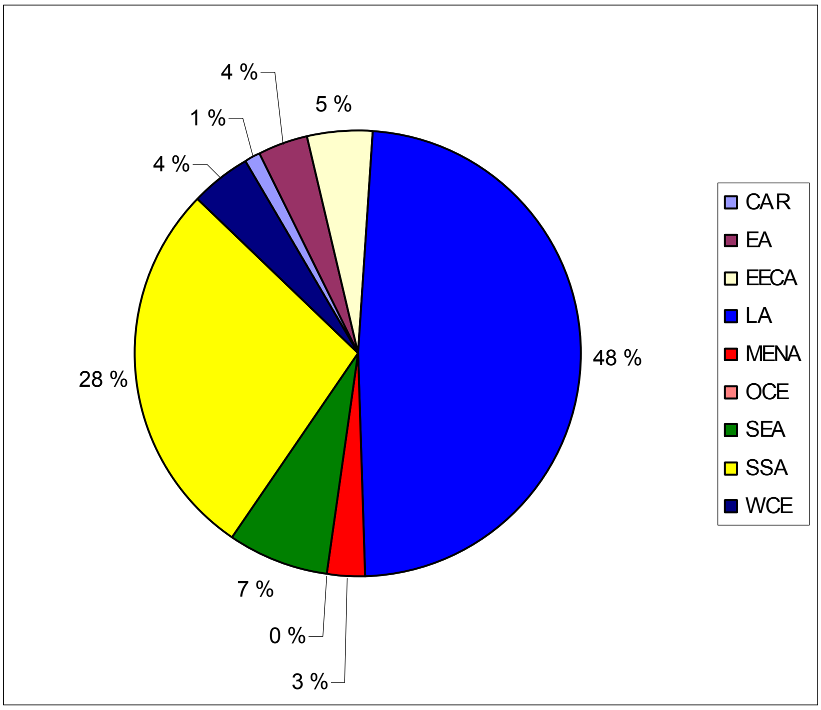
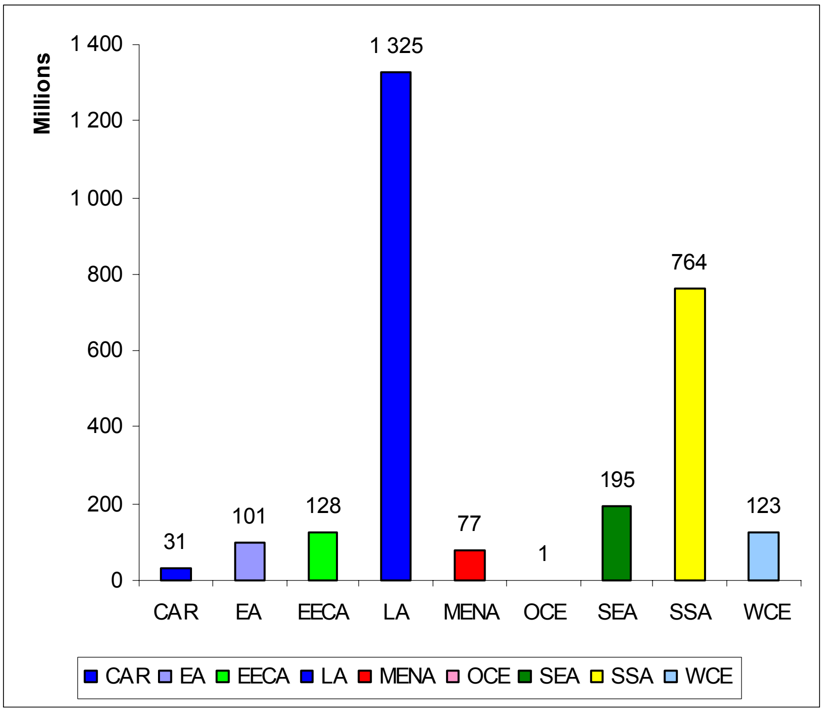


Figure 17 Total estimated expenditures for 136 low- and middle-income countries



The per capita spending gives a similar picture, but the Caribbean is now higher on the ranking. There are big variations between the regions; Latin America spends around 2.6 USD per capita on HIV and AIDS, while Eastern Asia only spends around ten cents per citizen on HIV and AIDS (figure 18). Overall, three regions have reached the one dollar per capita in AIDS expenditures using domestic public funds with the Caribbean being very close to this threshold.

Figure 18 Total estimated expenditures per capita for 136 low- and middle-income countries

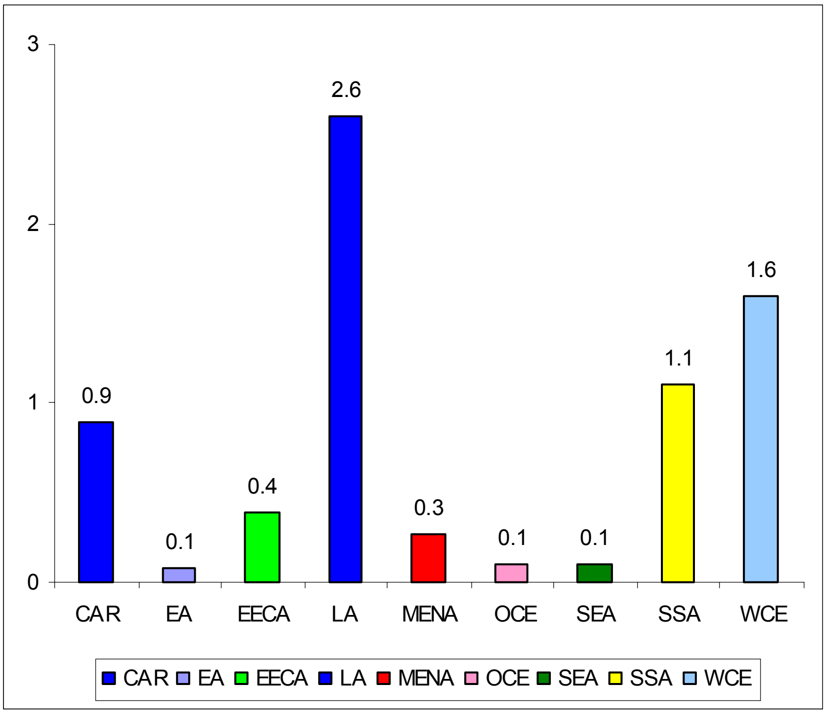


Figure 19 shows that in total, the lower middle-income countries are spending the largest amount on HIV and AIDS. However, when taking the size of the population into account, the upper middle-income countries has with the highest expenditures on HIV and AIDS (figure 20).

Figure 19 Estimated total expenditures for 136 countries by income level

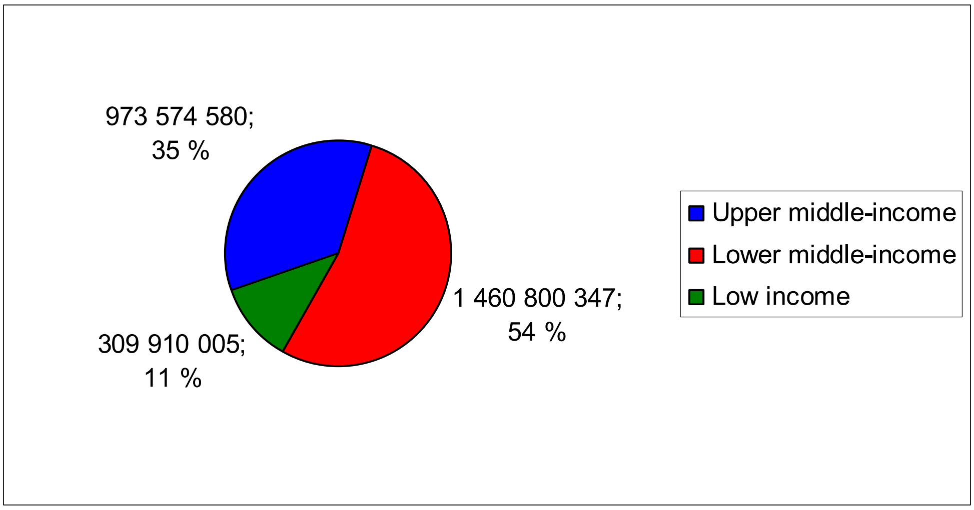


Figure 20 Estimated expenditures per capita for 136 countries by income level

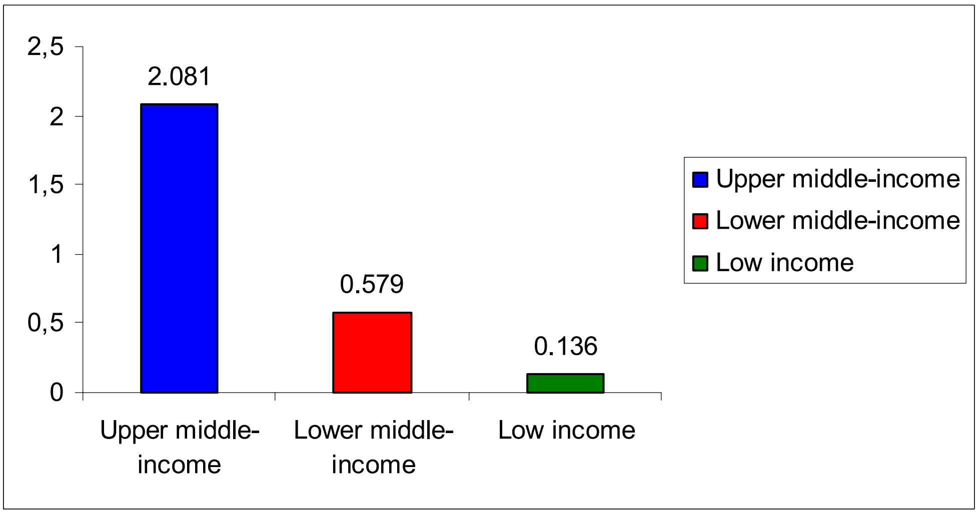
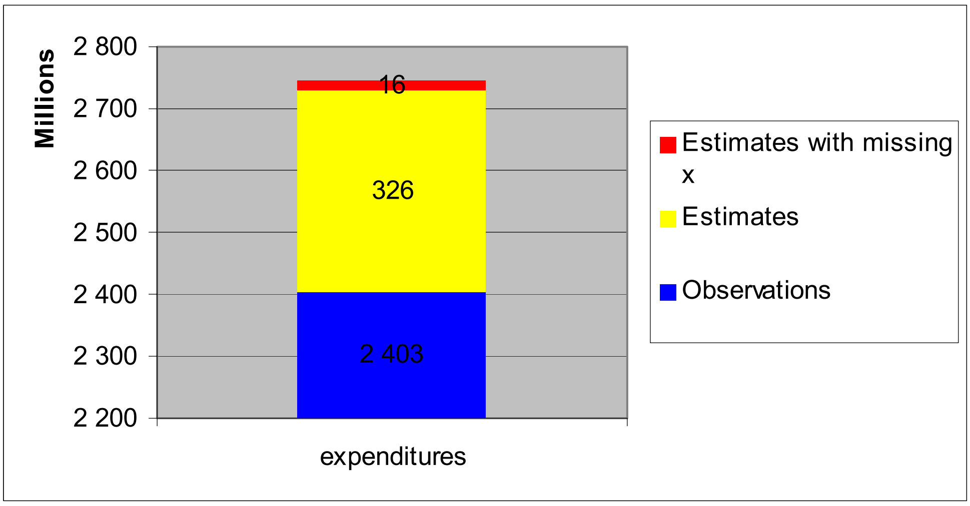


Figure 21 visualize how much the observed and estimated values are measuring in totals.

Figure 21 Total expenditures on HIV and AIDS by source of data



A measure often used in analysis of social security to give an indication of performance is the expenditures compared to a country's GDP. In this sense, Sub Saharan Africa is making the highest effort with about 0.15% of the GDP going to HIV and AIDS (figure 22). Eastern Asia Spends the least as percentage of GDP. From figure 23 we se that upper-middle income countries spends the most on HIV and AIDS as percentage of GDP, but the difference from what low income countries spend is only about 0.02%.

Figure 22 Estimated expenditures as % of GDP 131 low- and middle-income countries, by region

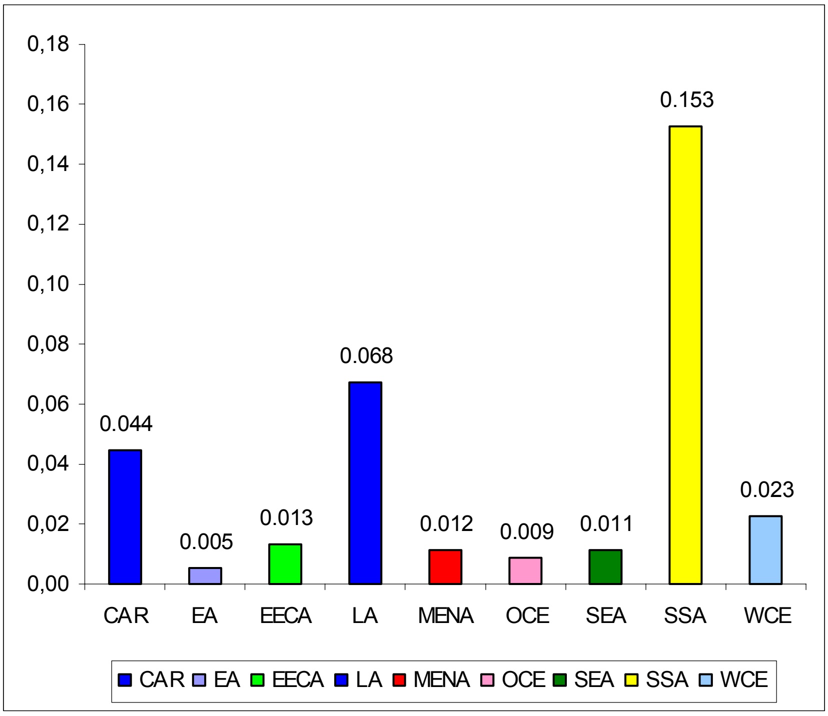
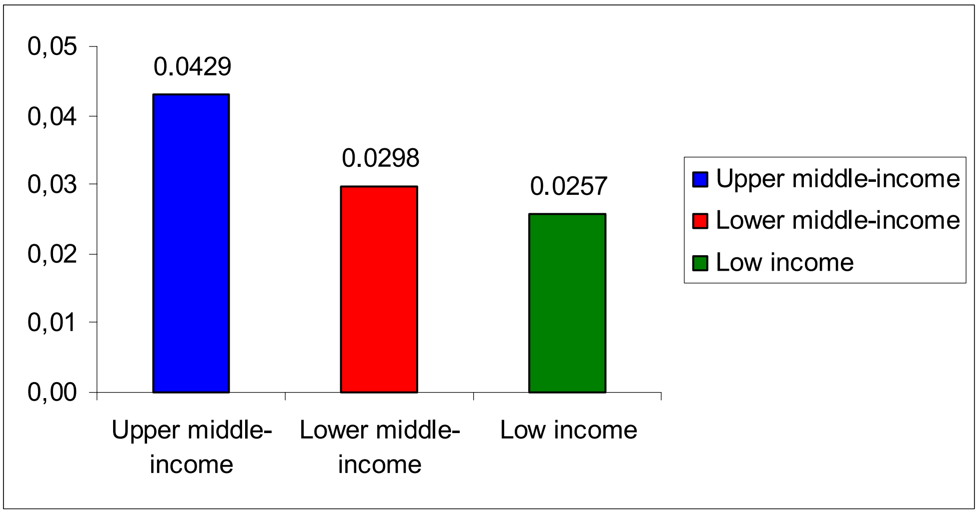


Figure 23 Estimated expenditures as % of GDP 131 low- and middle-income countries, by income level



6. Conclusion

In this paper we have presented a regression analysis based on a cross-section dataset for 81 countries to identify the main determinants of the level of domestic public expenditures on HIV and AIDS in low- and middle-income countries. The estimated model was used to predict the domestic public expenditures on HIV and AIDS for 2004 for the countries where observations on this variable were missing.

From the empirical analysis, four independent determinants for government spending was found; GDP per capita (coefficient estimate = 1.256; $p<0.001$), political stability (coefficient estimate = 0.587; $p=0.012$), HIV prevalence (coefficient estimate = 0.095; $p=0.002$), and budget support per capita (coefficient estimate = 0.059; $p=0.017$). There are regional variations for what determines spending in low- and middle-income countries. The level of income is the most influential determinant of the spending, as measured by the size of the coefficient in the model, but in the countries in Sub Saharan Africa the political stability is more important.

The main finding of the analysis is that combined observations and predictions for domestic public expenditures on HIV and AIDS in 2004 amounts to a total of 2.7 USD billions. The analysis revealed that there were regional differences in the level of expenditures; Latin America represents about 50% and Sub Saharan Africa represents about 30%. Taking the size of the population into account, Latin America still may be said to spend most, with 2.6 USD per capita, Sub Saharan Africa following with 1.1 USD. It is not unexpected that Latin America spends most since all but a few countries are middle-income countries. Furthermore, they have well-developed healthcare systems funded largely by governments (UNAIDS, 2006:234).

Eastern Asia and Europe has room for improvement when it comes to national effort to respond to the HIV pandemic. One striking finding was that Sub Saharan Africa comes out with the highest expenditures as % of GDP, and thus can be said to have made the best effort. Overall the governments in low- and middle-income countries spend 0.034% of their GDP on HIV and AIDS, and they only half a dollar per capita in 2004. However, when comparing the expenditures of regions, Purchase Power Parities should be applied in order to really capture that the relatively poorer countries might make a better effort than visible without these adjustments.

The aggregated domestic public expenditures for regions are plausible, and in line with earlier estimate (Dalen & Reijer, 2006). However, problems might arise when looking at each country specifically; Brazil and South Africa are spending more than predicted by the model. This could result from the fact that they spend a substantial amount on ART, specially the latest years, and ART is not accounted for in the model. Lithuania and other Eastern European countries are expected not to be as well explained by the model because only two observations from this region were available. When more information from this region becomes available, it should be included to improve the predictions for these countries.

When attempting to project the domestic public expenditures into the future, a problem that may arise is that the determinants as included as regressors in the multiple regression model could be difficult to assess or predict for future years when the estimated model based on the 2004 data is used for this purpose. An assumption might be that theses regressors either follow a certain trend estimated from available data, or simply that they remain constant.

Overall, based on our results, we may conclude that the analysis gives an estimation of the level of expenditures that should be presented aggregated by regional level. Furthermore, when no data is available on expenditures, this analysis provides, to the author's knowledge, the best estimate of domestic public expenditures on HIV and AIDS available today. When more information becomes available in the form of country-specific time series of the relevant variables, and when gaps in the data sets for some countries are filled, the approach presented her could be extended to a panel data analysis. In this way we could hope to give a better specification of unobserved heterogeneity across countries and to better understand the dynamics of the model.

7. References

Bulir A. & A. J. Hamann (2003): “Aid volatility: an empirical assessment”, IMF Staff Papers, 50: 64-89.

Dalen, H. P. & D. Reijer (2006): Projections of Funds for Population and AIDS activities 2005-2007, The Hague, Netherlands Interdisciplinary Demographic Institute, Resource Flows Project.

De Lay, P. R. Greener & J. A. Izazola (2007): “Are we spending too much on HIV? - No”, British Medical Journal, vol. 334, no. 7589:345.

England, R. (2007): “Are we spending too much on HIV? - Yes”. British Medical Journal, vol. 334, no. 7589:344.

The Global Fund to Fight AIDS Malaria and Tuberculoses (2007): “Guideline for proposals round 7”, Geneva, Switzerland.

The Global Fund to Fight AIDS Malaria and Tuberculoses (2003): “Guideline for proposals”, Geneva, Switzerland.

Hill, R.C., W. E. Griffiths & G. G. Judge (2001): Undergraduate Econometrics, John Wiley & Sons, Inc., New York, US.

Kaufmann, D., A. Kraay & M. Mastruzzi (2003): “Governance Matters III: Governance Indicators for 1996–2002”, World Bank Policy Research Working Paper 3106.

Lewis, M. (2005): “Addressing the Challenge of HIV/AIDS: Macroeconomic, Fiscal and Institutional Issues”, Centre for Global Development: Macroeconomic, Fiscal and Institutional Issues, Working Paper 58.

Menon-Johansson, A. S. (2005): “Good governance and good health: The role of societal structures in the human immunodeficiency virus pandemic”, BMC International Health and Human Rights 2005, 5:4.

Musgrove, P., R. Zeramdini, G. Carrin (2002): “Basic patterns in national health expenditure”, Bulletin of the World Health Organization, 80:134-142.

Scheffer, J. (2002): “Dealing with Missing Data”, Res. Lett. Inf. Math. Sci., 3: 153-160.

Secretary-General of the United Nations (2006): “Scaling up HIV prevention, treatment, care and support”, UNAIDS publication, Geneva, Switzerland.

Stata Base Reference Manual, release 9 (2005): [R] regress postestimation, Stata Press, vol. 3.

Stover, John (2007): Presentation at the UNAIDS Advisory Board meeting, Glion, Switzerland.

Sydsæter, K., A. Strøm & P. Berck (1999): Economists Mathematical Manual, (third ed.), Springer, Berlin, Germany.

Tayler, L. & C. Dickinson (2005): "The link between corruption and HIV/AIDS", Global Corruption Report 2006, ch. 6, Transparency international, Berlin, Germany.

Transparency International (2004): Sources for the Transparency International Corruption Perceptions Index 2004, Transparency International, Berlin, Germany.

UNAIDS (2007): "2007 United Nations General Assembly Special Session on HIV/AIDS – Guidelines on construction of core indicators", UNAIDS, Geneva Switzerland.

UNAIDS (2006): "2006 Report on the global AIDS epidemic", UNAIDS Geneva Switzerland.

UNAIDS (2005): "Resource needs for an expanded response to AIDS in low- and middle-income countries", Geneva, Switzerland.

UNGASS (2006): "UNGASS 2005 reporting: Debriefing on the UNGASS reporting process and recommendations for process & content improvement", Geneva, Switzerland.

WHO (2002): "To the fifty-second session of the Regional Committee for Africa, Brazzaville, Republic of Congo, 8 to 12 October 2002 – The work of WHO in the African region 2000-2001".

World Bank (2006): "A Decade of Measuring the Quality of Governance", Governance Matters 2006, Washington, US.

Xu, K, D.B. Evans, K. Kawabata, R. Zeramdini, J. Klavus & C.J.L. Murray (2003): "Household catastrophic health expenditure: A multicountry analysis", Lancet, 362: 111–117.

8. Annex

Table A1 Variables with number of data points when including the latest year available

Main variable	Proxy Variable	Countries covered (2004/2004 and latest available)
Amount of money available	GDP/cap	131
	GNI/cap	131
	Total debt service (% of GNI)	124
	Tax revenue (% of GDP)	50/76
Ability to spend the money	Number of hospital beds (per 10 000 population)	41/93
	Nurses (density per 1 000 population)	54/136
	Literacy rate	138
	Physicians (density per 1 000 population)	52/137
	Political stability	137
	Corruption Index	108
Commitment/will	Per capita government expenditure on health at average exchange rate (US$)	138
	Government expenditure on health at average exchange rate (US$) as % of total health expenditures	136
	Per capita government expenditure on health at international dollar rate	138
	Per capita total expenditure on health at average exchange rate (US$)	138
	ODA Commitments in 2004 (2004 in constant dollars) as share of domestic public health expenditures	135
	ODA Commitments in 2004 (2004 in constant dollars) as share of people living with HIV	120
	Number of NGOs in the country	0
	Strategic plan	0
	2005 possible composite index from UCC survey	~89
Burden of the disease	Estimated number of people living with HIV, children (0-14)	115
	Estimated number of people living with HIV, adult and children	127

	Estimated number of people living with HIV, adults (15+)	126
	HIV prevalence, adult (14-49)	126
	Estimated HIV prevalence, children (0-14)	113
	Number of people with advanced HIV infection (urgent need)	138 (113)
	People with advanced HIV infection receiving antiretroviral (ARV) combination therapy (%)	124
Dependent variable	Domestic public expenditures related to HIV/AIDS	65/81

Table A2 The correlation matrix shows that political stability has highest correlation with HIV expenditures, and that the variables are collinear

obs=79							
	Indom	Political Stability	Voice and accountability	Government effectiveness	Regulatory quality	Rule of law	Control of corruption
Indomex	1						
Political Stability	0.5571	1					
Voice and accountability	0.5275	0.6658	1				
Government effectiveness	0.5179	0.6452	0.7999	1			
Regulatory quality	0.4668	0.6348	0.8082	0.8725	1		
Rule of law	0.4792	0.8033	0.8298	0.8929	0.8465	1	
Control of corruption	0.5271	0.7116	0.8071	0.8829	0.8323	0.9242	1

Table A3 Correlation matrix between the domestic public expenditures on HIV and AIDS and the potential explanatory variables, number of observations in the row below

	lndomex	domexcap	gdpcap	budgetcap	debtshare	taxrev	beds	nurses	physic	literacy	polstab	cpi
GDP /cap	0.5237	0.2524	1									
	80	80	131									
Budget support /cap	0.1093	-0.0074	0.1087	1								
	81	81	131	138								
Total debt service (% of GNI)	0.1781	-0.141	0.2279	-0.0035	1							
	78	78	123	124	124							
Tax revenue (% of GDP)	0.1624	0.2676	0.1822	0.1531	0.0076	1						
	46	46	74	76	72	76						
Number of hospital beds (per 10 000 population)	0.1919	0.1809	0.357	0.0535	0.0933	0.4244	1					
	50	50	87	93	82	59	93					
Nurses (per 1 000 population)	0.1588	0.0557	0.4811	-0.0498	0.269	0.1668	0.7957	1				
	80	80	130	136	123	75	92	136				
Physicians (density per 1 000 population)	0.0875	-0.0499	0.4309	-0.0263	0.2739	-0.0068	0.6962	0.7341	1			
	81	81	131	137	124	76	93	136	137			
Literacy rate	-0.0111	0.2309	-0.017	-0.0784	0.1116	0.0417	0.1409	0.2718	0.2515	1		
	81	81	131	138	124	76	93	136	137	138		
Political stability	0.5634	0.2579	0.6063	0.0944	0.1505	0.4009	0.36	0.3197	0.2631	-0.0071	1	
	80	80	130	137	123	75	93	135	136	137	137	
Corruption Index	-0.0062	-0.0046	-0.0125	-0.0432	0.0297	-0.0331	-0.1338	-0.0711	-0.0855	0.1339	-0.0104	1
	61	61	102	108	96	58	74	107	107	108	107	108
Government expenditure on health at average exchange rate (US$) as % of	0.3601	0.0927	0.4526	0.155	0.1166	0.3869	0.4332	0.3782	0.2461	0.0268	0.476	-0.1393

Total health expenditures												
	81	81	131	136	123	76	92	135	136	136	135	106
ODA Commitments for HIV and AIDS in 2004 as a ratio of domestic public health expenditures	-0.2609	-0.0528	-0.2489	0.029	-0.1333	-0.1766	-0.0528	-0.2052	-0.2832	-0.2132	-0.1344	-0.0845
	81	81	131	135	123	76	91	134	135	135	134	106
ODA Commitments for HIV and AIDS in 2004 divided by people living with HIV	-0.2515	-0.0447	-0.0653	0.059	-0.0524	-0.1892	-0.0601	-0.038	-0.0528	-0.0143	0.013	0.0083
	75	75	115	120	113	69	77	119	120	120	119	92
People with advanced HIV infection receiving antiretroviral combination therapy (%)	0.5336	0.3356	0.4558	-0.0634	0.2336	0.0022	-0.1308	0.0329	0.201	0.0709	0.3487	0.0752
	74	74	117	122	113	68	80	121	122	122	121	95
Estimated number of people living with HIV, children (0-14)	0.0016	-0.0139	-0.1998	0.1413	-0.1025	0.0131	-0.1497	-0.1538	-0.2676	-0.1787	-0.1774	0.0379
	73	73	110	115	107	64	72	114	115	115	114	90
HIV prevalence, children (0-14)	0.2566	0.3109	-0.1614	0.1123	-0.187	0.3384	-0.2159	-0.1065	-0.3418	-0.0473	0.0161	-0.0298
	73	73	110	113	107	63	70	112	113	113	112	89
Estimated number of people living with HIV, adult and children	0.0262	0.0095	-0.1124	0.0353	-0.0244	0.0037	-0.0946	-0.0887	-0.1712	-0.0818	-0.1274	0.1031
	76	76	121	127	118	73	84	126	127	127	126	98
Estimated number of people living with HIV, adults (15+)	0.032	0.0101	-0.104	0.0272	-0.0165	0.0017	-0.091	-0.0811	-0.1616	-0.0719	-0.1205	0.104

	75	75	120	126	117	72	83	125	126	126	125	97
HIV prevalence, adult (14-49)	0.3444	0.4468	-0.0979	0.0723	-0.1619	0.3453	-0.1526	-0.0695	-0.3107	0.0165	0.0995	0.0204
	76	76	121	127	118	73	84	126	127	127	126	98
number of people in need of ART	0.0688	0.0134	-0.1014	0.0403	-0.0236	0.0076	-0.0966	-0.0761	-0.1699	-0.099	-0.1153	0.0653
	73	73	108	113	106	63	70	112	113	113	112	88

Table A3 Continued

	hlthshare	odashare	odaplwh	ART	child	child prev	plwh	adult	prevalence	needART
Government expenditure on health at average exchange rate (US$) as % of Total health expenditures	1									
	136									
ODA Commitments for HIV and AIDS in 2004 as a ratio of domestic public health expenditures	-0.11	1								
	135	135								
ODA Commitments for HIV and AIDS in 2004 as a ratio of people living with HIV	0.20	0.10	1							
	119	118	120							
People with advanced HIV infection receiving antiretroviral combination therapy (%)	0.16	-0.13	-0.01	1						
	121	120	113	122						
Estimated number of people living with HIV, children (0-14)	-0.21	0.31	-0.08	-0.20	1					
	114	113	114	108	115					
HIV prevalence, children (0-14)	0.03	0.25	-0.07	-0.06	0.52	1				
	112	112	112	107	113	113				
Estimated number of people living with HIV, adult and children	-0.22	0.11	-0.07	-0.14	0.82	0.30	1			
	126	125	120	120	115	113	127			
Estimated number of people living with HIV, adults (15+)	-0.22	0.09	-0.07	-0.14	0.80	0.28	1.00	1		
	125	124	119	119	115	113	126	126		
HIV prevalence, adult (14-49)	-0.01	0.16	-0.09	0.01	0.46	0.95	0.35	0.33	1	
	126	125	119	119	115	113	126	126	127	
number of people in need of ART	-0.19	0.06	-0.07	-0.14	0.83	0.34	0.98	0.98	0.35	1
	112	111	113	106	113	111	113	113	113	113

Figure A1 The scatter plot reveals that the untransformed GDP/cap has a concave relationship with the transformed expenditures rather than a linear one

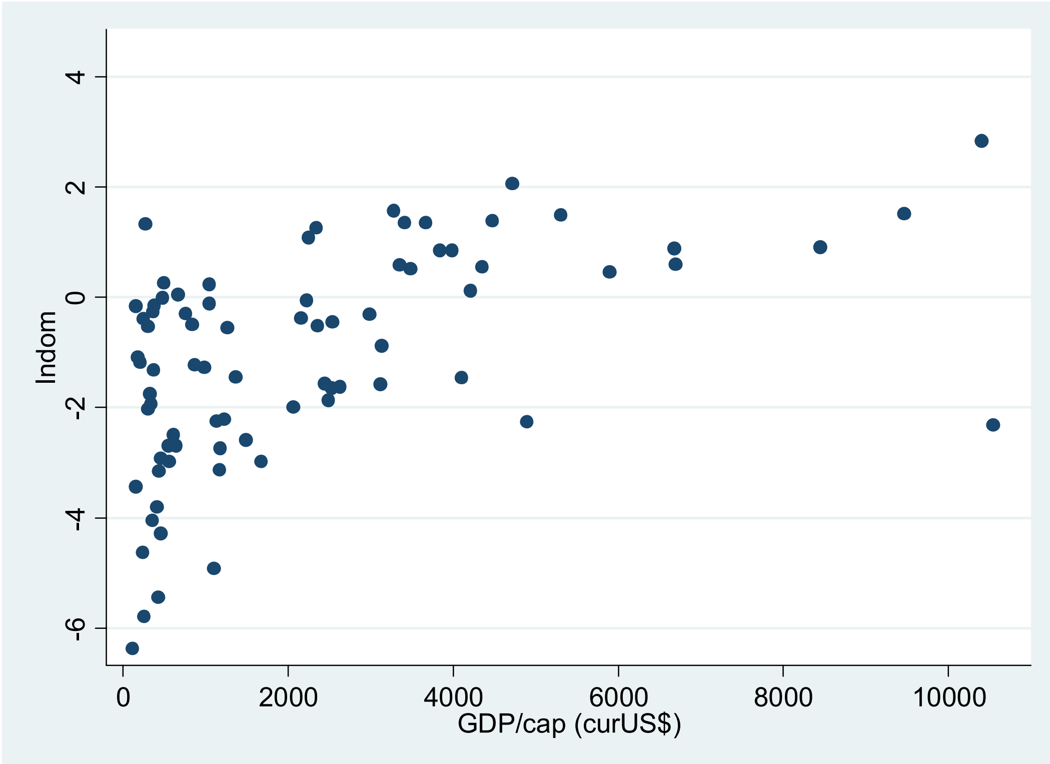


Figure A2 The augmented component-plus-residual plot confirms the concave relationship of the GDP/cap and the transformed expenditures

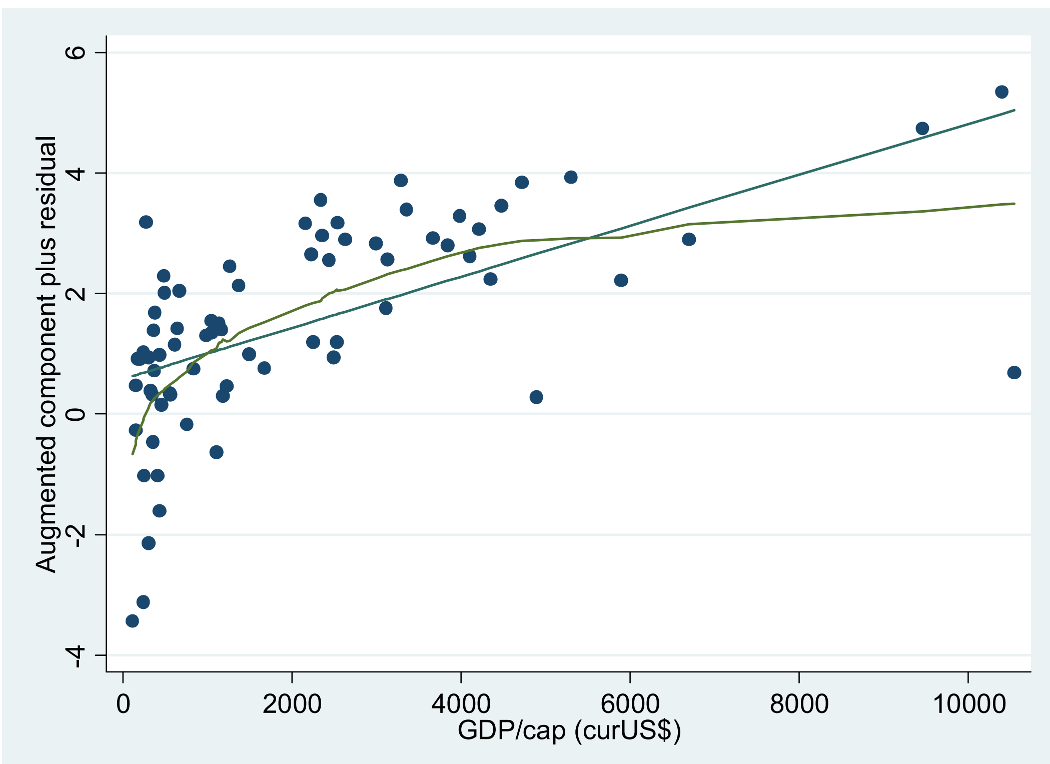


Table A4 The forward regression with the significant variables

	lndomex	lndomex	lndomex	lndomex
Political stability	1.367***	0.898***	0.812**	0.736**
	(0.227)	(0.252)	(0.252)	(0.251)
logarithm of GDP	--	0.634***	0.749***	0.848***
		(0.178)	(0.168)	(0.173)
HIV prevalence adult	--	--	0.109***	0.111***
15-49 years			(0.026)	(0.026)
Budget support per c	--	--	--	0.052#
apita				(0.027)
Intercept	-4.647***	-7.941***	-8.884***	-9.564***
	(0.630)	(1.095)	(1.052)	(1.094)
N	80	79	74	74
R2	0.317	0.415	0.525	0.549
Adjusted R2	0.309	0.400	0.505	0.522
*** p<0.001; ** p<0.01; * p<0.05; # p<0.1; two tailed				

Table A5 Regression results and joint hypothesis testing of the six dimensions of governance

```
      Source |       SS       df       MS              Number of obs =      74
-------------+------------------------------           F(  9,    64) =   10.40
       Model |  182.782388     9  20.3091543           Prob > F      =  0.0000
    Residual |  124.935704    64  1.95212038           R-squared     =  0.5940
-------------+------------------------------           Adj R-squared =  0.5369
       Total |  307.718093    73  4.21531634           Root MSE      =  1.3972

------------------------------------------------------------------------------
     lndomex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |   .6801085   .2313035     2.94   0.005     .2180266     1.14219
   budgetcap |   .0282713   .0310812     0.91   0.366    -.0338206    .0903632
  prevalence |   .1108772    .026416     4.20   0.000     .0581052    .1636493
    p2olstab |   .8588105   .3611302     2.38   0.020       .13737    1.580251
       Voice |   .7644409   .4067491     1.88   0.065    -.0481338    1.577016
  gvrnmnteff |  -.1235399   .8927554    -0.14   0.890    -1.907024    1.659944
  regquality |   .1453756   .5431219     0.27   0.790    -.9396351    1.230386
   ruleoflaw |  -1.715947   1.077919    -1.59   0.116    -3.869337    .4374424
  corruption |   1.121959   .7835789     1.43   0.157    -.4434202    2.687337
       _cons |  -8.782887   1.818829    -4.83   0.000    -12.41642   -5.149359
------------------------------------------------------------------------------

 ( 1)  p2olstab = 0
 ( 2)  Voice = 0
 ( 3)  gvrnmnteff = 0
 ( 4)  regquality = 0
 ( 5)  ruleoflaw = 0
 ( 6)  corruption = 0

       F(  6,    64) =    2.67
```

```
           Prob > F =    0.0223

 ( 1)  Voice = 0
 ( 2)  gvrnmnteff = 0
 ( 3)  regquality = 0
 ( 4)  ruleoflaw = 0
 ( 5)  corruption = 0

       F(  5,    64) =    1.43
            Prob > F =    0.2246
```

Figure A3 The residual plot of the model without dummy variables

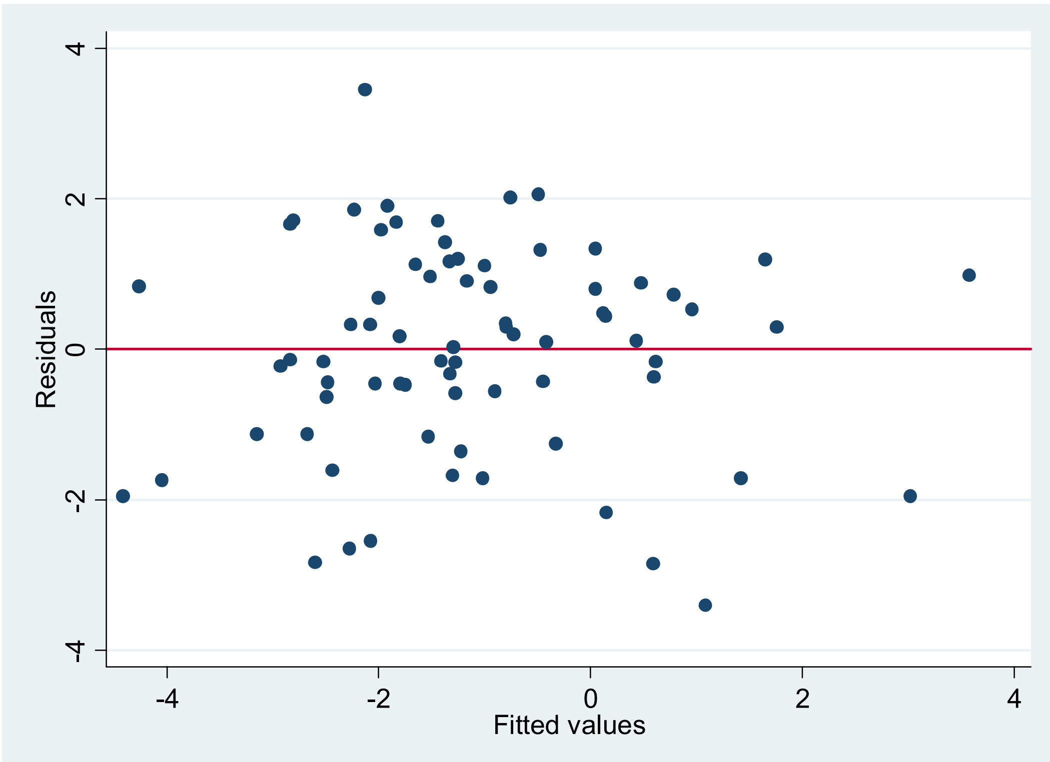


Table A6 The 10 lowest and highest observations on the residual before including dummy variables

e	region
-3.405611	WCE
-2.854148	SSA
-2.834659	SEA
-2.64726	SEA
-2.55023	SSA
-2.178003	SSA
-1.953101	SSA
-1.948666	SSA
-1.738438	SEA
-1.718504	EECA

```
+-------------------+
|        e   region |
|-------------------|
| 1.584975      SSA |
| 1.657287      SSA |
| 1.682247      SSA |
| 1.696752      SSA |
| 1.711589      SSA |
|-------------------|
| 1.851469       LA |
| 1.907939      SSA |
| 2.012213       LA |
| 2.053335       LA |
| 3.449041      SSA |
+-------------------+
```

Table A7 Stem-and-leaf plot for the residual of the model without dummy variables

```
e rounded to nearest multiple of .01
plot in units of .01

 -3** | 41
 -2** | 85,83,65,55
 -2** | 18
 -1** | 95,95,74,72,72,68,62
 -1** | 36,26,16,13,13
 -0** | 63,59,56
 -0** | 47,46,46,44,43,37,33,22,18,17,17,16,14
  0** | 03,10,11,17,20,29,29,33,33,34,44,47
  0** | 53,68,73,80,83,83,87,90,96,98
  1** | 11,12,16,19,20,32,33,42
  1** | 58,66,68,70,71,85,91
  2** | 01,05
  2** |
  3** | 45
```

Table A8 Inclusion of the different dummies, coefficients and standard errors in parenthesis

	lndomex	lndomex	lndomex	lndomex	lndomex
Political stability	0.611*	0.758**	0.600*	0.591*	0.587*
	(0.251)	(0.238)	(0.231)	(0.243)	(0.227)
logarithm of GDP	1.122***	0.667***	0.988***	1.474***	1.256***
	(0.210)	(0.175)	(0.197)	(0.250)	(0.244)
HIV prevalence adult	0.072*	0.127***	0.079**	0.095**	0.095**
15-49 years	(0.031)	(0.025)	(0.029)	(0.031)	(0.029)
Budget support per c	0.058*	0.052*	0.061*	0.056*	0.059*
apita	(0.027)	(0.026)	(0.025)	(0.026)	(0.024)
Sub-Saharan Africa	1.094*	--	1.417**	6.400**	5.197*

	(0.501)		(0.470)	(2.255)	(2.148)
Latin America	--	1.232**	1.463***	--	1.301**
		(0.419)	(0.404)		(0.407)
reg8gdp	--	--	--	-0.799*	-0.575#
				(0.332)	(0.319)
Intercept	-11.488***	-8.683***	-11.009***	-14.110***	-12.948***
	(1.383)	(1.081)	(1.280)	(1.723)	(1.656)
N	74	74	74	74	74
R2	0.578	0.599	0.647	0.612	0.664
Adjusted R2	0.547	0.570	0.616	0.577	0.628
*** p<0.001; ** p<0.01; * p<0.05; # p<0.1; two tailed					

Figure A4 The logarithm of GDP per capita for Sub-Saharan Africa, and without Sub-Saharan Africa

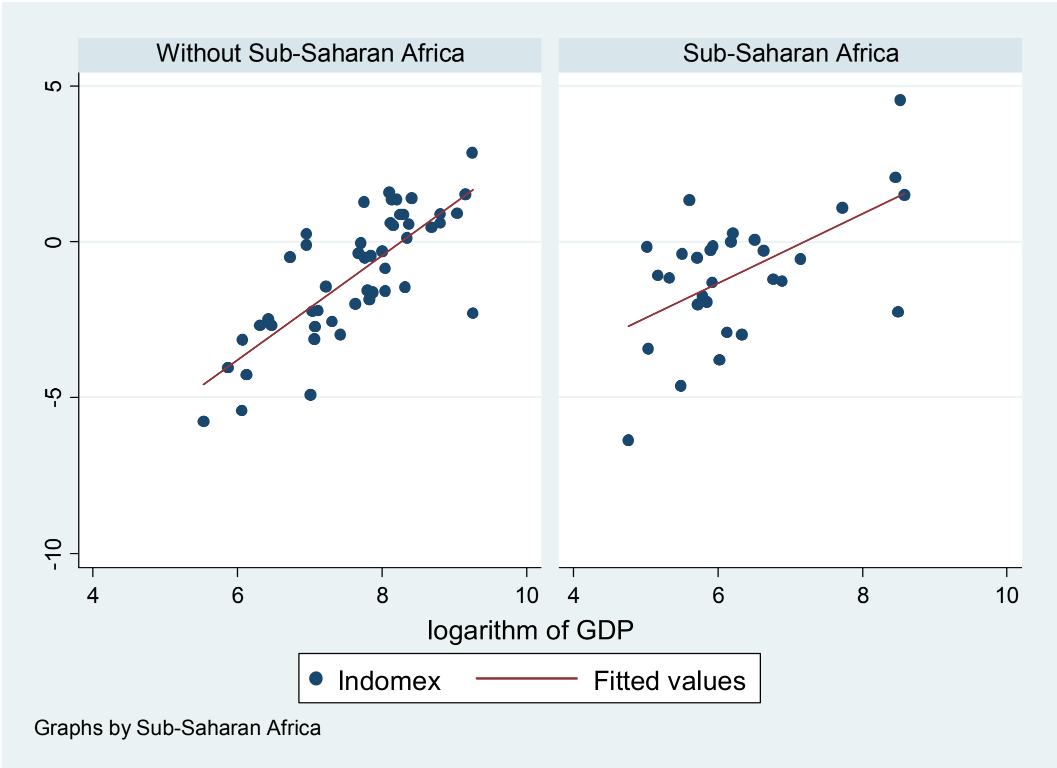


Table A9 Stem-and-leaf plot for estud (Studentized residuals)

```
estud rounded to nearest multiple of .01
plot in units of .01
```

```
-2** | 97
-2** | 68,67
-2** | 55
-2** | 21
-2** |
-1** |
-1** |
-1** | 53,43
-1** | 39,35,31
-1** | 13
-0** | 92,88
-0** | 77,76,75,75
-0** | 54,54,53,49,42
-0** | 39,20
-0** | 17,15,09,07,05,01
 0** | 04,04,07,07,10,11,18
 0** | 23,26,29,32,36,36,37
 0** | 40,40,47,54,55,55,56
 0** | 60,60,68,70,71,72,73,76,78,78,79
 0** | 83,84,89
 1** | 02,04,06,10,14
 1** | 22,30
 1** | 52
 1** |
 1** |
 2** |
 2** |
 2** |
 2** | 75
```

Table A10 The 10 lowest and highest observations on studentized residuals

studentized residuals	countries
-2.967592	Czech Republic
-2.681854	Madagascar
-2.665214	Mauritius
-2.548771	Congo, democratic
-2.209551	Mozambique
-1.529407	Philippines
-1.429628	Swaziland
-1.38774	Lesotho
-1.347651	Laos
-1.314407	Guinea, Republic
.8886309	Senegal
1.018214	Guinea-Bissau
1.038743	Burkina Faso
1.055346	Brazil
1.103635	Benin

```
|------------------------------|
| 1.139333         El Salvador |
| 1.221211               Kenya |
| 1.304484            Barbados |
| 1.518213            Botswana |
| 2.750422              Gambia |
+------------------------------+
```

Table A11 Stem-and-leaf plot for leverage

```
leverage rounded to nearest multiple of .001
plot in units of .001

  0** | 41,42,44,44,44,45,46,46,47,48,48,49,51,53,53,56,56,56,56,56,57
  0** | 60,60,61,61,61,61,62,62,64,64,65,65,65,66,67,67,68,69,70,71, ... (30)
  0** | 80,81,81,81,82,83,83,84,84,87,88,89,90,91,91,92,93,94,94,95, ... (28)
  1** | 00,00,02,09,15,16,16
  1** | 20,21,22,23,24,24,27,28,29,38,39
  1** | 40,42,44,44,47,51,52,54,55
  1** | 63,63,64,65,69,79
  1** | 85,87,91
  2** |
  2** | 20
  2** |
  2** | 68
  2** |
  3** | 09
  3** | 26
  3** |
  3** |
  3** |
  4** |
  4** |
  4** |
  4** | 77
```

Table A12 The five highest observations on leverage

```
+---------------------+
| leverage  countries |
|---------------------|
| .2202888      Gabon |
| .2683715   Botswana |
| .3085434  Mauritius |
| .3256772  Swaziland |
| .4769459     Guyana |
+---------------------+
```

Figure A5 The standardized normal probability plot

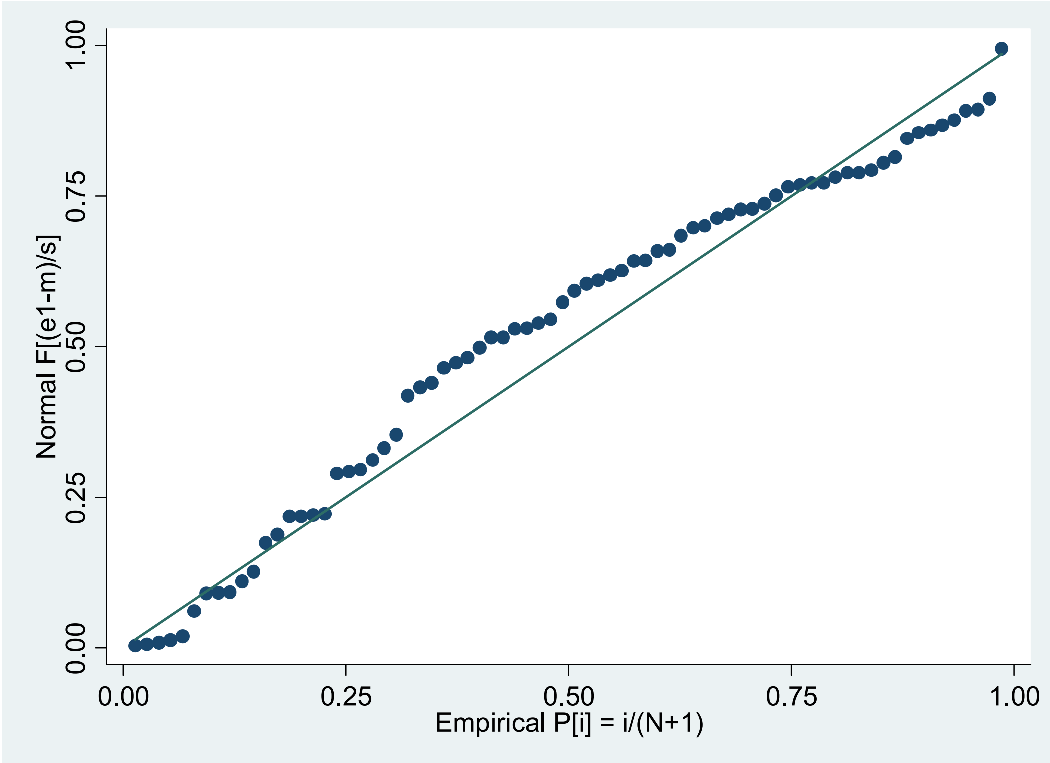


Figure A6 The quantiles of the residuals plotted with the quantiles of the normal distribution

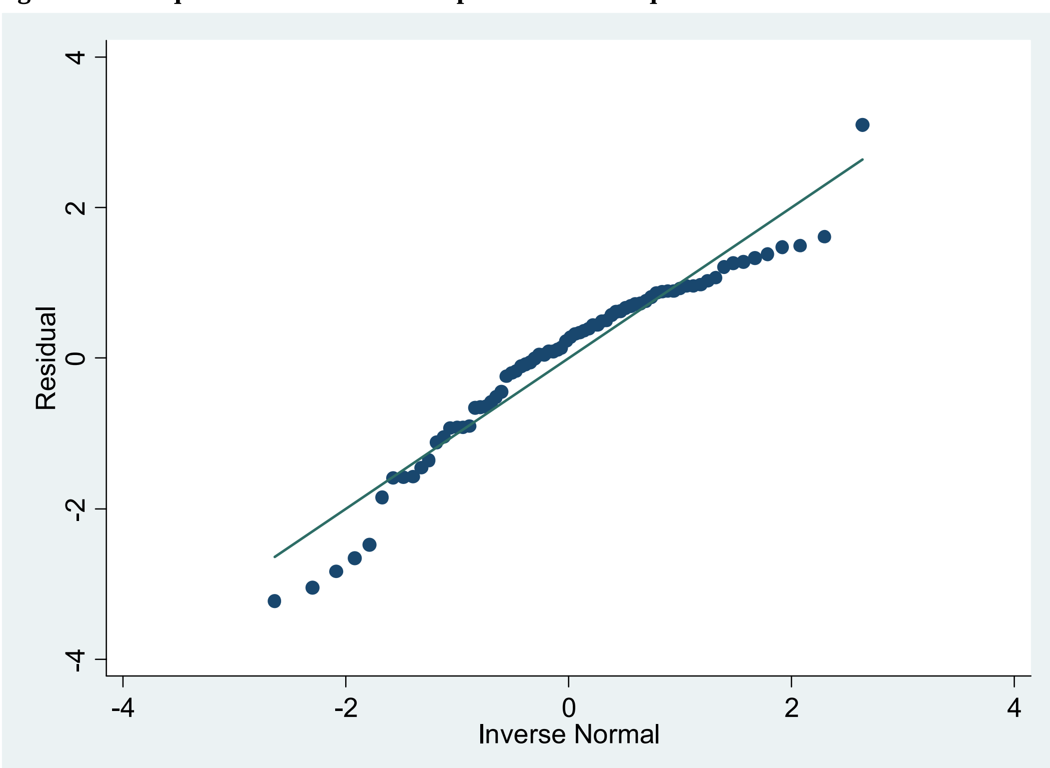


Table A13 White test for Heteroskedasticity, the null hypothesis being that the residuals are heterogeneous

Source	chi2	degrees of freedom	p-value
Heteroskedasticity	42.13	28	0.0421

Table A14 Breusch-Pagan / Cook-Weisberg test for heteroskedasticity, , the null hypothesis being that the residuals are heterogeneous

```
Ho: Constant variance
Variables: fitted values of lndomex

chi2(1)      =     0.78
Prob > chi2  =   0.3768
```